

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:34:52 ; Search time 55.3333 Seconds  
(without alignments)  
64.831 Million cell updates/sec

Title: US-10-008-377a-2

Perfect score: 52  
Sequence: 1 GVLLWEIFSL 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 444336

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :  
1: Genesep23Sep04:\*  
2: Genesep1980s:\*  
3: Genesep2000s:\*  
4: Genesep2001s:\*  
5: Genesep2002s:\*  
6: Genesep2003as:\*  
7: Genesep2003bs:\*  
8: Genesep2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	10	6	AAO22984
2	46	88.5	9	5	AAE31409
3	46	88.5	9	6	AAO22987
4	46	88.5	9	7	ADK65216
5	46	88.5	9	7	ADK65225
6	46	88.5	9	7	ADK65275
7	46	88.5	9	7	ADK65285
8	46	88.5	9	7	ADK65266
9	46	88.5	9	7	ADK65235
10	46	88.5	9	8	ADM97287
11	42	80.8	10	8	ADM97302
12	32	61.5	10	4	AAQ34960
13	31	59.6	9	2	AAW37279
14	31	59.6	10	3	AAW37279
15	28	53.8	9	2	AAW37279
16	28	53.8	9	2	AAW37279
17	28	53.8	9	2	AAW37279
18	28	53.8	9	2	AAW37279
19	28	53.8	9	2	AAW37279
20	28	53.8	9	5	ABP30946
21	28	53.8	9	7	ADK65216
22	28	53.8	9	7	ADK65216
23	28	53.8	9	7	ADK65216
24	28	53.8	9	8	ADM97287
25	28	53.8	9	8	ADM97287

26	28	53.8	10	2	AAV46028
27	28	53.8	10	2	AAV46431
28	28	53.8	10	2	AAV46511
29	28	53.8	10	4	AAU02657
30	28	53.8	10	5	ABP30929
31	28	53.8	10	5	ABP30941
32	28	53.8	10	7	ADA08574
33	28	53.8	10	7	ADA08586
34	28	53.8	10	7	ADP08929
35	28	53.8	10	7	ADP08917
36	28	53.8	10	7	ADG46204
37	28	53.8	10	7	ADG46216
38	27	51.9	8	3	AAV84415
39	27	51.9	9	2	AAV84415
40	27	51.9	10	7	ADL17633
41	26	50.0	8	4	AAV84415
42	26	50.0	8	5	ADG03135
43	26	50.0	9	4	AAV84415
44	26	50.0	9	4	AAV84415
45	26	50.0	9	6	AAV84415

## ALIGNMENTS

RESULT 1	AAO22984	standard; peptide; 10 AA.
ID	AAO22984	
XX	AAO22984	
AC	AAO22984	
XX	17-SEP-2003	(first entry)
DT	17-SEP-2003	
XX	Human p376-85 ALK-derived HLA-A*0201 restricted CTL epitope peptide.	
DE	Human p376-85 ALK-derived HLA-A*0201 restricted CTL epitope peptide.	
XX	HLA-A*0201 binding; anaplastic lymphoma kinase; ALK; CTL epitope; ALCL;	
KW	cytotoxic T-lymphocyte; receptor tyrosine; NPM/ALK fusion protein; human;	
KW	cytotoxic; anaplastic large cell lymphoma; neuroblastoma; gene therapy;	
KW	t(2;5) translocation; neoplasia; cytostatic; immunotherapy; alkoma;	
KW	p376-85.	
OS	Human sapiens.	
XX	WO2003042243-A2.	
PN	WO2003042243-A2.	
XX	22-MAY-2003.	
PD	22-MAY-2003.	
XX	14-NOV-2002; 2002WO-EP012764.	
PF	14-NOV-2002; 2002WO-EP012764.	
XX	15-NOV-2001; 2001US-00008377.	
PR	15-NOV-2001; 2001US-00008377.	
XX	(NAST-) 1ST NAZ STUDIO E CURA DEI TUMORI.	
PA	(STUD-) ST JUDE CHILDREN'S RES HOSPITAL.	
XX	Gambacorti-Passerini C, Passoni L;	
PI	Gambacorti-Passerini C, Passoni L;	
XX	WPI; 2003-441791/41.	
DR	WPI; 2003-441791/41.	
XX	New HLA-A*0201-binding Anaplastic Lymphoma Kinase (ALK) peptide,	
PT	useful for preparing a composition for treating ALK-positive lymphoma,	
PT	neuroblastoma or ALK-expressing neoplasia.	
XX	Claim 1; Page 7; 33pp; English.	
PS	Claim 1; Page 7; 33pp; English.	
XX	The invention relates to a novel HLA-A*0201-binding anaplastic lymphoma	
CC	kinase (ALK)-derived peptide, of 9-10 amino acids, which may act as a	
CC	cytotoxic T-lymphocyte (CTL)-epitope. ALK is a receptor tyrosine kinase	
CC	which forms part of the NPM/ALK fusion protein (oncogenic lymphoma kinase	
CC	fusion protein) in ALCL (anaplastic large cell lymphoma)-derived cell	
CC	lines. More than 50% of ALCL cases possess a t(2;5) chromosomal	
CC	translocation that leads to the expression of the NPM/ALK fusion protein	
CC	which forms a potent oncogene when constitutively activated. Translocated	
CC	ALK is a widely expressed tumour-associated antigen characteristic of ALK	

-positive lymphomas, neuroblastomas and ALK-expressing neoplasias. The peptides of the invention demonstrate cytostatic activity and induce an MHC (major histocompatibility complex) class I restricted cytotoxic lymphocyte response against tumour cells expressing the NPM/ALK fusion protein. Hence, the peptides may be utilised during the treatment, via immunotherapy, of ALK-positive lymphomas such as alkomas, neuroblastomas and ALK-expressing neoplasias. In addition, the peptides may be used during gene therapy. The current sequence is that of the human p376-85 ALK-derived HLA-A\*0201 restricted CTL epitope peptide of the invention

XX  
SQ Sequence 10 AA;

Query Match 100.0%; Score 52; DB 6; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLLMEIFSL 10  
Db 1 GVLLMEIFSL 10

RESULT 2  
AAE31409  
ID AAE31409 standard; peptide; 9 AA.  
XX  
AC AAE31409;

DT 24-FEB-2003 (first entry)

DE Human VGR3 peptide #3.

XX Human; T-lymphocyte; vaccine; viral infection; gene therapy; cancer.

XX Homo sapiens.

XX WO200272627-A2.

XX 19-SEP-2002.

XX 11-MAR-2002; 2002WO-EP002666.

XX 09-MAR-2001; 2001US-0274250P.

XX 14-MAY-2001; 2001US-0290353P.

XX 18-MAY-2001; 2001US-0291610P.

XX (CALL-) CALLISTOGEN AG.

XX Wrede F, Walden P, Eichler-Wertens M, Filter W,

XX WPI; 2002-759836/82.

XX Providing, identifying or optimizing peptides for inducing cytotoxic T-lymphocytes and for treating cancer, comprises selecting conserved PT regions in antigenic proteins and identifying CD8+ T-cell epitopes in the protein.

XX Disclosure; Page 15; 32pp; English.

XX The invention relates to a method for providing, identifying or/and CC optimizing peptides which induce cytotoxic T-lymphocytes and to the uses CC of the obtained peptides for vaccination. The method is useful for CC providing, identifying and/or optimizing peptides that are useful in CC manufacturing a pharmaceutical composition for the induction of cytotoxic CC T-lymphocytes, and for the prevention, treatment or diagnosis of cancer CC or viral infections. The invention is also used in gene therapy. The CC present sequence is human VGR3 peptide used to illustrate the method of CC the invention

XX Sequence 9 AA;

Query Match 88.5%; Score 46; DB 5; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLLMEIFSL 10  
Db 1 VLLMEIFSL 9

RESULT 3  
AAO22987  
ID AAO22987 standard; peptide; 9 AA.  
XX  
AC AAO22987;

DT 17-SEP-2003 (first entry)

DE Human p377-85 ALK-derived HLA-A\*0201 restricted CTL epitope peptide.

XX HLA-A\*0201 binding; anaplastic lymphoma kinase; ALK; CTL epitope; ALCL;  
XX cytotoxic T-lymphocyte; receptor tyrosine; NPM/ALK fusion protein; human;  
XX oncogenic; anaplastic large cell lymphoma; neuroblastoma; gene therapy;  
XX t(2;5) translocation; neoplasia; cytostatic; immunotherapy; alkoma;  
XX p377-85.

XX Homo sapiens.

XX WO2003042243-A2.

XX 22-MAY-2003.

XX 14-NOV-2002; 2002WO-EP012764.

XX 15-NOV-2001; 2001US-0008377.

XX (MAST-) ISR NAZ STUDIO E CURA DEI TUMORI.

XX (SUDD-) ST JUDE CHILDREN'S RES HOSPITAL.

XX Gambacorti-Passerini C, Passoni L;

XX WPI; 2003-441791/41.

XX New HLA-A\*0201-binding Anaplastic Lymphoma Kinase (ALK) peptide,  
PT useful for preparing a composition for treating ALK-positive lymphoma,  
PT neuroblastoma or ALK-expressing neoplasia.

XX Claim 1; Page 7; 33pp; English.

XX The invention relates to a novel HLA-A\*0201-binding anaplastic lymphoma CC kinase (ALK)-derived peptide, of 9-10 amino acids, which may act as a CC cytotoxic T-lymphocyte (CTL)-epitope. ALK is a receptor tyrosine kinase CC which forms part of the NPM/ALK fusion protein (oncogenic lymphoma kinase CC fusion protein) in ALCL (anaplastic large cell lymphoma)-derived cell CC lines. More than 50% of ALCL cases possess a t(2;5) chromosomal CC translocation that leads to the expression of a t(2;5) ALK fusion protein CC which forms a potent oncogene when constitutively activated. Translocated CC ALK is a widely expressed tumour-associated antigen characteristic of ALK CC -positive lymphomas, neuroblastomas and ALK-expressing neoplasias. The CC peptides of the invention demonstrate cytostatic activity and induce an CC MHC (major histocompatibility complex) class I restricted cytotoxic CC lymphocyte response against tumour cells expressing the NPM/ALK fusion CC protein. Hence, the peptides may be utilised during the treatment, via CC immunotherapy, of ALK-positive lymphomas such as alkomas, neuroblastomas CC and ALK-expressing neoplasias. In addition, the peptides may be used CC during gene therapy. The current sequence is that of the human p377-85 CC ALK-derived HLA-A\*0201 restricted CTL epitope peptide of the invention

XX Sequence 9 AA;

Query Match 88.5%; Score 46; DB 6; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLLMEIFSL 10  
Db 1 VLLMEIFSL 9

RESULT 4  
ADK65216  
ID ADK65216 standard; peptide; 9 AA.  
XX  
XX  
AC ADK65216;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Human VEGFR-1-derived peptide #2 for anti-angiogenesis treatment.

XX antiangiogenic; cytoskeletal; antiinflammatory; immunosuppressive;  
KW ophthalmological; gynecological; antiarteriosclerotic; virucide;  
KW hepatocytic; dermatological; anti-HIV; antidiabetic; antipsoriatic;  
KW antirheumatic; antiarthritic; antithyroid; immunization; angiogenesis;  
KW vascular endothelial growth factor; VEGF; neuropilin;  
KW placental growth factor; tumor; neoplasias; metastases; inflammation;  
KW autoimmunity; eye disease; arthritis; endometriosis; arteriosclerosis;  
KW edema; hepatitis; Kaposi sarcoma; diabetes; psoriasis;  
KW rheumatoid arthritis; thyroiditis; diabetic retinopathy;  
KW transplant rejection; macular degeneration; neovascular glaucoma;  
KW hemangioma; angiofibroma.  
XX  
OS Homo sapiens.  
XX  
PN WO2003086450-A1.  
XX  
PD 23-OCT-2003.  
XX  
PF 11-APR-2003; 2003WO-CU000004.  
XX  
PR 15-APR-2002; 2002CU-00000076.  
XX  
PA (INGG-) CENT ING GENETICA & BIOTECNOLOGIA.  
XX  
PI Bequet Romero M, Acevedo Castro BE, Gavilondo Cowley JV;  
PI Fernandez Molina LE, Lopez Ochoa O, Silva Rodriguez RDC;  
PI Musachio Lasa A, Galban Rodriguez E, Vazquez Blomquist DM;  
XX  
DR WPI; 2003-833615/77.

XX Active immunization against angiogenic proteins, useful for treating e.g.  
PT tumors and inflammation, particularly contains vascular endothelial  
PT growth factor or its receptor.  
XX  
PS Disclosure; Page 19; 53pp; Spanish.

XX The invention relates to an active immunization against angiogenic  
CC proteins comprising administration of a vaccination composition (A),  
CC optionally containing an adjuvant, that comprises polypeptides (I)  
CC directly associated with an increase in angiogenesis, their variants, or  
CC their encoding polynucleotides (II). Angiogenesis-associated polypeptides  
CC are: members of the vascular endothelial growth factor (VEGF) family,  
CC especially the 121, 165 or 189 isoforms of VEGF-A, the 167 isoform of  
CC VEGF-B, or VEGF-C or -D; a (co-)receptor of VEGF, particularly VEGFR-1,  
CC 2 or -3; NRP-1 or -2 (members of the neuropilin family); or placental  
CC growth factor. (A) is used for treatment or prevention of tumors in  
CC mammals, particularly humans but also farm animals and pets, also many  
CC other conditions associated with excessive angiogenesis, specifically  
CC malignant or benign neoplasias (and their metastases), acute or chronic  
CC inflammation, autoimmunity and eye diseases (claimed). Among the diseases  
CC that may be treated are arthritis, endometriosis, arteriosclerosis,  
CC edema, infectious diseases (hepatitis and Kaposi sarcoma), diabetes,  
CC psoriasis, rheumatoid arthritis, thyroiditis, diabetic retinopathy,  
CC transplant rejection, macular degeneration, neovascular glaucoma,  
CC hemangioma and angiofibroma. The method destroys cells that are the  
CC source of angiogenic proteins, rather than just neutralizing the activity  
CC of such proteins (as in passive immunization). This sequence represents  
CC an immunisation peptides of the invention derived from the VEGF receptor  
XX proteins.  
XX  
SQ Sequence 9 AA;

Query Match 88.5%; Score 46; DB 7; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLLMEIFSL 10  
|||||||  
Db 1 VLLMEIFSL 9

RESULT 5  
ADK65225  
ID ADK65225 standard; peptide; 9 AA.  
XX  
XX  
AC ADK65225;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Human VEGFR-2-derived peptide #1 for anti-angiogenesis treatment.

XX antiangiogenic; cytoskeletal; antiinflammatory; immunosuppressive;  
KW ophthalmological; gynecological; antiarteriosclerotic; virucide;  
KW hepatocytic; dermatological; anti-HIV; antidiabetic; antipsoriatic;  
KW antirheumatic; antiarthritic; antithyroid; immunization; angiogenesis;  
KW vascular endothelial growth factor; VEGF; neuropilin;  
KW placental growth factor; tumor; neoplasias; metastases; inflammation;  
KW autoimmunity; eye disease; arthritis; endometriosis; arteriosclerosis;  
KW edema; hepatitis; Kaposi sarcoma; diabetes; psoriasis;  
KW rheumatoid arthritis; thyroiditis; diabetic retinopathy;  
KW transplant rejection; macular degeneration; neovascular glaucoma;  
KW hemangioma; angiofibroma.  
XX  
OS Homo sapiens.  
XX  
PN WO2003086450-A1.  
XX  
PD 23-OCT-2003.  
XX  
PF 11-APR-2003; 2003WO-CU000004.  
XX  
PR 15-APR-2002; 2002CU-00000076.  
XX  
PA (INGG-) CENT ING GENETICA & BIOTECNOLOGIA.  
XX  
PI Bequet Romero M, Acevedo Castro BE, Gavilondo Cowley JV;  
PI Fernandez Molina LE, Lopez Ochoa O, Silva Rodriguez RDC;  
PI Musachio Lasa A, Galban Rodriguez E, Vazquez Blomquist DM;  
XX  
DR WPI; 2003-833615/77.

XX Active immunization against angiogenic proteins, useful for treating e.g.  
PT tumors and inflammation, particularly contains vascular endothelial  
PT growth factor or its receptor.  
XX  
PS Disclosure; Page 19; 53pp; Spanish.

XX The invention relates to an active immunization against angiogenic  
CC proteins comprising administration of a vaccination composition (A),  
CC optionally containing an adjuvant, that comprises polypeptides (I)  
CC directly associated with an increase in angiogenesis, their variants, or  
CC their encoding polynucleotides (II). Angiogenesis-associated polypeptides  
CC are: members of the vascular endothelial growth factor (VEGF) family,  
CC especially the 121, 165 or 189 isoforms of VEGF-A, the 167 isoform of  
CC VEGF-B, or VEGF-C or -D; a (co-)receptor of VEGF, particularly VEGFR-1,  
CC 2 or -3; NRP-1 or -2 (members of the neuropilin family); or placental  
CC growth factor. (A) is used for treatment or prevention of tumors in  
CC mammals, particularly humans but also farm animals and pets, also many  
CC other conditions associated with excessive angiogenesis, specifically  
CC malignant or benign neoplasias (and their metastases), acute or chronic  
CC inflammation, autoimmunity and eye diseases (claimed). Among the diseases  
CC that may be treated are arthritis, endometriosis, arteriosclerosis,  
CC edema, infectious diseases (hepatitis and Kaposi sarcoma), diabetes,  
CC psoriasis, rheumatoid arthritis, thyroiditis, diabetic retinopathy,  
XX

CC transplant rejection, macular degeneration, neovascular glaucoma,  
 CC hemangioma and angiofibroma. The method destroys cells that are the  
 CC source of angiogenic proteins, rather than just neutralizing the activity  
 CC of such proteins (as in passive immunization). This sequence represents  
 CC an immunisation peptides of the invention derived from the VEGF receptor  
 CC proteins.  
 XX  
 XX Sequence 9 AA;  
 SQ  
 Query Match 88.5%; Score 46; DB 7; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 VLWEIFSL 10  
 Db 1 VLWEIFSL 9  
 RESULT 6  
 ADK65275  
 ID ADK65275 standard; peptide; 9 AA.  
 XX  
 AC ADK65275;  
 XX  
 DT 06-MAY-2004 (first entry)  
 DE  
 XX Human VEGFR-2-derived peptide #11 for anti-angiogenesis treatment.  
 XX  
 XX antiangiogenic; cytostatic; antiinflammatory; immunosuppressive;  
 KW ophthalmological; gynecological; antiarteriosclerotic; virucide;  
 KW hepatotropic; dermatological; anti-HIV; antidiabetic; antipsoriatic;  
 KW antirheumatic; antiarthritic; antithyroid; immunization; angiogenesis;  
 KW vascular endothelial growth factor; VEGF; neuropilin;  
 KW placental growth factor; tumor; neoplasias; metastases; inflammation;  
 KW autoimmunity; eye disease; arthritis; endometriosis; arteriosclerosis;  
 KW edema; hepatitis; Kaposi sarcoma; diabetes; psoriasis;  
 KW rheumatoid arthritis; thyroiditis; diabetic retinopathy;  
 KW rheumatoid arthritis; macular degeneration; neovascular glaucoma;  
 KW transplant rejection; angiofibroma.  
 KW hemangioma; angiofibroma.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO2003086450-A1.  
 PN  
 XX 23-OCT-2003.  
 PD  
 XX 11-APR-2003; 2003WO-CU000004.  
 PF  
 XX 15-APR-2002; 2002CU-00000076.  
 PR  
 XX (INGG-) CENT ING GENETICA & BIOTECNOLOGIA.  
 PA  
 XX Bequet Romero M, Acevedo Castro BE, Gavilondo Cowley JV,  
 PI Fernandez Molina LE, Lopez Oceljo O, Silva Rodriguez RDC;  
 PI Musachio Lasa A, Galban Rodriguez E, Vazquez Blomquist DM;  
 XX  
 XX WPI; 2003-833615/77.  
 DR  
 XX  
 XX Active immunization against angiogenic proteins, useful for treating e.g.  
 PT tumors and inflammation, particularly contains vascular endothelial  
 PT growth factor or its receptor.  
 PT  
 XX Disclosure; Page 19; 53pp; Spanish.  
 PS  
 XX The invention relates to an active immunization against angiogenic  
 CC proteins comprising administration of a vaccination composition (A),  
 CC optionally containing an adjuvant, that comprises polypeptides (I),  
 CC directly associated with an increase in angiogenesis, their variants, or  
 CC their encoding polynucleotides (II). Angiogenesis-associated polypeptides  
 CC are: members of the vascular endothelial growth factor (VEGF) family,  
 CC especially the 121, 165 or 189 isoforms of VEGF-A, the 167 isoform of  
 CC VEGF-B, or VEGF-C or -D; a (co-)receptor of VEGF, particularly VEGFR-1,  
 CC 2 or -3, NRP-1 or -2 (members of the neuropilin family); or placental

CC growth factor. (A) is used for treatment or prevention of tumors in  
 CC mammals, particularly humans but also farm animals and pets, also many  
 CC other conditions associated with excessive angiogenesis, specifically  
 CC malignant or benign neoplasias (and their metastases), acute or chronic  
 CC inflammation, autoimmunity and eye diseases (claimed). Among the diseases  
 CC that may be treated are arthritis, endometriosis, arteriosclerosis,  
 CC edema, infectious diseases (hepatitis and Kaposi sarcoma), diabetes,  
 CC psoriasis, rheumatoid arthritis, thyroiditis, diabetic retinopathy,  
 CC transplant rejection, macular degeneration, neovascular glaucoma,  
 CC hemangioma and angiofibroma. The method destroys cells that are the  
 CC source of angiogenic proteins, rather than just neutralizing the activity  
 CC of such proteins (as in passive immunization). This sequence represents  
 CC an immunisation peptides of the invention derived from the VEGF receptor  
 CC proteins.  
 XX  
 XX Sequence 9 AA;  
 SQ

Query Match 88.5%; Score 46; DB 7; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VLWEIFSL 10  
 Db 1 VLWEIFSL 9

RESULT 7  
 ADK65285  
 ID ADK65285 standard; peptide; 9 AA.  
 XX  
 AC ADK65285;  
 XX  
 DT 06-MAY-2004 (first entry)  
 DE

XX Human VEGFR-3-derived peptide #11 for anti-angiogenesis treatment.

XX antiangiogenic; cytostatic; antiinflammatory; immunosuppressive;  
 KW ophthalmological; gynecological; antiarteriosclerotic; virucide;  
 KW hepatotropic; dermatological; anti-HIV; antidiabetic; antipsoriatic;  
 KW antirheumatic; antiarthritic; antithyroid; immunization; angiogenesis;  
 KW vascular endothelial growth factor; VEGF; neuropilin;  
 KW placental growth factor; tumor; neoplasias; metastases; inflammation;  
 KW autoimmunity; eye disease; arthritis; endometriosis; arteriosclerosis;  
 KW edema; hepatitis; Kaposi sarcoma; diabetes; psoriasis;  
 KW rheumatoid arthritis; thyroiditis; diabetic retinopathy;  
 KW transplant rejection; macular degeneration; neovascular glaucoma;  
 KW hemangioma; angiofibroma.

XX Homo sapiens.  
 OS  
 XX WO2003086450-A1.  
 PN  
 XX 23-OCT-2003.  
 PD  
 XX 11-APR-2003; 2003WO-CU000004.  
 PF  
 XX 15-APR-2002; 2002CU-00000076.  
 PR  
 XX (INGG-) CENT ING GENETICA & BIOTECNOLOGIA.  
 PA

XX Bequet Romero M, Acevedo Castro BE, Gavilondo Cowley JV,  
 PI Fernandez Molina LE, Lopez Oceljo O, Silva Rodriguez RDC;  
 PI Musachio Lasa A, Galban Rodriguez E, Vazquez Blomquist DM;  
 XX  
 XX WPI; 2003-833615/77.  
 DR

XX Active immunization against angiogenic proteins, useful for treating e.g.  
 PT tumors and inflammation, particularly contains vascular endothelial  
 PT growth factor or its receptor.

XX Disclosure; Page 19; 53pp; Spanish.  
 PS  
 XX The invention relates to an active immunization against angiogenic

CC	proteins comprising administration of a vaccination composition (A),
CC	optionally containing an adjuvant, that comprises polypeptides (I),
CC	directly associated with an increase in angiogenesis, their variants, or
CC	their encoding polynucleotides (II). Angiogenesis-associated polypeptides
CC	are: members of the vascular endothelial growth factor (VEGF) family,
CC	especially the 121, 165 or 189 isoforms of VEGF-A, the 167 isoform of
CC	VEGF-B, or VEGF-C or -D; a (co-)receptor of VEGF, particularly VEGFR-1,
CC	2 or -3; NRP-1 or -2 (members of the neuropilin family); or placental
CC	growth factor. (A) is used for treatment or prevention of tumors in
CC	mammals, particularly humans but also farm animals and pets, also many
CC	other conditions associated with excessive angiogenesis, specifically
CC	malignant or benign neoplasias (and their metastases), acute or chronic
CC	inflammation, autoimmunity and eye diseases (claimed). Among the diseases
CC	that may be treated are arthritis, endometriosis, arteriosclerosis,
CC	edema, infectious diseases (hepatitis and Kaposi sarcoma), diabetes,
CC	psoriasis, rheumatoid arthritis, thyroiditis, diabetic retinopathy,
CC	transplant rejection, macular degeneration, neurovascular glaucoma,
CC	hemangioma and angiodysplasia. The method destroys cells that are the
CC	source of angiogenic proteins, rather than just neutralizing the activity
CC	of such proteins (as in passive immunization). This sequence represents
CC	an immunisation peptides of the invention derived from the VEGF receptor
CC	proteins.
CC	
XX	
SQ	Sequence 9 AA;
Query Match	89.5%; Score 46; DB 7; Length 9;
Best Local Similarity	100.0%; Pred. No. 1.7e+06;
Matches	9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
2 VLMEFSL 10	
1 VLMEFSL 9	
RESULT 8	
ADK65266	standard; peptide; 9 AA.
XX	
ADK65266;	
XX	
06-MAY-2004 (first entry)	
DT	
XX	
Human VEGFR-1-derived peptide #12 for anti-angiogenesis treatment.	
XX	
KW	antiangiogenic; cytostatic; antiinflammatory; immunosuppressive;
KW	ophthalmological; gynecological; antiartherosclerotic; vitreous;
KW	heparotrophic; dermatological; anti-HIV; antidabetic; antipsoriatic;
KW	antithrombotic; antiarthritic; antihypertensive; immunization; angiogenesis;
KW	vascular endothelial growth factor; VEGF; neuropilin;
KW	placental growth factor; tumor; neoplasias; metastases; inflammation;
KW	autoimmunity; eye disease; arthritis; endometriosis; arteriosclerosis;
KW	edema; hepatitis; Kaposi sarcoma; diabetes; psoriasis;
KW	rheumatoid arthritis; thyroiditis; diabetic retinopathy;
KW	transplant rejection; macular degeneration; neovascular glaucoma;
KW	hemangioma; angiodysplasia.
XX	
OS	Homo sapiens.
XX	
PN	WO2003086450-A1.
XX	
PD	23-OCT-2003.
XX	
PF	11-APR-2003; 2003WO-CU000004.
XX	
PR	15-APR-2002; 2002CU-00000076.
(INGG-) CENT ING GENETICA & BIOTECNOLOGIA.	
PA	Bequet Romero M, Acevedo Castro BE, Gavilondo Cowley JV;
PI	Fernandez Molina LE, Lopez Ochoa O, Silva Rodriguez RDL;
PI	Musachio Laaa A, Galban Rodriguez E, Vazquez Blomquist DM;
WP:	2003-833615/77

XX PT Active immunization against angiogenic proteins, useful for treating e.g.  
PT tumors and inflammation, particularly contains vascular endothelial  
XX growth factor or its receptor.

PS Disclosure; Page 19; 53pp; Spanish.

CC The invention relates to an active immunisation against angiogenic  
CC proteins comprising administration of a vaccination composition (A),  
CC optionally containing an adjuvant, that comprises polypeptides (I),  
CC directionally associated with an increase in angiogenesis, their variants, or  
CC directly associated with an increase in angiogenesis, their variants, or  
CC their encoding polynucleotides (II). Angiogenesis-associated polypeptides  
CC are: members of the vascular endothelial growth factor (VEGF) family,  
CC especially the 121, 165 or 189 isoforms of VEGF-A, the 167 isoform of  
CC VEGF-B, or VEGF-C or -D, a (co-)receptor of VEGF, particularly VEGFR-1, -  
CC 2 or -3, NRP-1 or -2 (members of the neuropilin family); or placental  
CC growth factor. (A) is used for treatment or prevention of tumors in  
CC mammals, particularly humans but also farm animals and pets, also many  
CC other conditions associated with excessive angiogenesis, specifically  
CC malignant or benign neoplasias (and their metastases), acute or chronic  
CC inflammation, autoimmunity and eye diseases (cataract). Among the diseases  
CC that may be treated are arthritis, endometriosis, arteriosclerosis,  
CC edema, infectious diseases (hepatitis and Kaposi's sarcoma), diabetes,  
CC psoriasis, rheumatoid arthritis, thyroiditis, diabetic retinopathy,  
CC transplacental rejection, macular degeneration, neovascular glaucoma,  
CC hemangioma, angiodysplasia.

SQ Sequence 9 AA:

Query Match Best Local Similarity 88.5%; Score 46; DB 7; Length 9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 VLMEHIFSL 10  
| | | | | | | | | |  
Db 1 VLMEHIFSL 9

RESULT\_9  
ADK65235  
ID ADK65235 standard; peptide; 9 AA.

XX ADK65235;  
XX  
DT 06-MAY-2004 (first entry)  
DE Human VEGFR-3-derived peptide #1 for anti-angiogenesis treatment.  
XX  
XX antiangiogenic; cytostatic; antiinflammatory; immunosuppressive;  
XX ophthalmological; gynecological; antidiabetic; antiviral;  
XX hepatotropic; dermatological; anti-HIV; antidyslipidemic; antiproliferative;  
XX antineoplastic; antirheumatic; antithyroid; immunization; angiogenesis;  
XX vascular endothelial growth factor; VEGF; neuropilin;  
XX placental growth factor; tumor; neoplasia; metastasis; inflammation;  
XX autoimmunity; eye disease; arthritis; endometriosis; arteriosclerosis;  
XX edema; hepatitis; Kaposi's sarcoma; diabetes; psoriasis;  
XX rheumatoid arthritis; thyroiditis; diabetic retinopathy;  
XX transplant rejection; macular degeneration; neovascular glaucoma;  
XX hemangioma; angiodysplasia.

CS Homo sapiens.  
XX  
XX W02003086450-A1.  
XX  
XX 23-OCT-2003.  
XX  
XX 11-APR-2003; 2003WO-CU000004.  
XX  
XX 15-APR-2002; 2002GT-000000076.  
XX

XX (ING-) CENT ING GENETICA & BIOTECNOLOGIA.

XX Beget Romero M, Acevedo Castro BE, Gavilondo Cowley JV;

XX Fernandez Molina LE, Lopez Ocoje O, Silva Rodriguez RDC;

XX Musachio Lasa A, Galban Rodriguez E, Vazquez Blomquist DM;

XX WPI; 2003-833615/77.

XX Active immunization against angiogenic proteins, useful for treating e.g. tumors and inflammation, particularly contains vascular endothelial growth factor or its receptor.

XX Disclosure; Page 19; 53pp; Spanish.

XX The invention relates to an active immunization against angiogenic proteins comprising administration of a vaccination composition (A), optionally containing an adjuvant, that comprises polypeptides (I) directly associated with an increase in angiogenesis, their variants, or their encoding polynucleotides (II). Angiogenesis-associated polypeptides are: members of the vascular endothelial growth factor (VEGF) family, especially the 121, 165 or 189 isoforms of VEGF-A, the 167 isoform of VEGF-B, or VEGF-C or -D, a (co-)receptor of VEGF, particularly VEGFR-1, 2 or -3; NRP-1 or -2 (members of the neuropilin family); or placental growth factor. (A) is used for treatment or prevention of tumors in mammals, particularly humans but also farm animals and pets, also many other conditions associated with excessive angiogenesis, specifically malignant or benign neoplasias (and their metastases), acute or chronic inflammation, autoimmunity and eye diseases (cataracts, arteriosclerosis, edema, infectious diseases (hepatitis and Kaposi sarcoma), diabetes, psoriasis, rheumatoid arthritis, thyroiditis, diabetic retinopathy, transplants rejection, macular degeneration, neovascular glaucoma, hemangioma and angiodysplasia). The method destroys cells that are the source of angiogenic proteins, rather than just neutralizing the activity of such proteins (as in passive immunization). This sequence represents an immunisation peptide of the invention derived from the VEGF receptor proteins.

XX Sequence 9 AA;

Query Match 88.5%; Score 46; DB 7; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLMEIFSL 10  
1 VLMEIFSL 9

Db

RESULT 10  
ADM97287  
ADM97287 standard; peptide; 9 AA.

AC ADM97287;

DT 17-JUN-2004 (first entry)

DE Cytotoxic T cell inducing KDR peptide SEQ ID NO: 26.

KW vaccine; cytostatic; cancer; antiarthritic; antirheumatic; antidiabetic;

KW antiarteriosclerotic; ophthalmological; cytotoxic T cell; CTL; VEGFR-2;

KW vascular endothelial growth factor receptor 2; KDR peptide.

OS Homo sapiens.

PN WO2004024766-A1.

PD 25-MAR-2004.

PF 12-SEP-2003; 2003WO-JP011722.

PR 12-SEP-2002; 2002JP-00267285.

PR 07-MAR-2003; 2003JP-00062003.

PR 11-JUN-2003; 2003JP-00167042.

PA (ONCO-) ONCOTHERAPY SCI INC.

PA (TAHA/) TAHARA H.

XX Tahara H, Wada S, Tsunoda T;

XX WPI; 2004-295074/27.

XX Nonapeptides comprising cytotoxic T-cell inducing activity, useful as vaccine for treating proliferation of malignant tumors.

XX Example 6; Page 82; 100pp; Japanese.

XX The present invention relates to nona- and decapeptides comprising cytotoxic T-cell (CTL) inducing activity which are KDR peptides binding to HLA-A2402, containing partial sequences of receptor protein VEGFR-2 (Vascular endothelial growth factor receptor 2), or derived from these sequences by addition or substitution of one or more amino acid residues. The peptides can be used to produce a cancer vaccine for treatment and prevention of cancer (including stomach, duodenum, colon, lung, breast, prostate and brain cancer) and suppression of proliferation and metastasis of malignant tumors. The vaccine is also useful for treatment of diabetic retinopathy, chronic rheumatoid arthritis and atherosclerosis. The present sequence is a peptide of the invention.

XX Sequence 9 AA;

Query Match 88.5%; Score 46; DB 8; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLMEIFSL 10  
1 VLMEIFSL 9

Db

RESULT 11  
ADM97302  
ADM97302 standard; peptide; 10 AA.

AC ADM97302;

DT 17-JUN-2004 (first entry)

DE Cytotoxic T cell inducing KDR peptide SEQ ID NO: 41.

KW vaccine; cytostatic; cancer; antiarthritic; antirheumatic; antidiabetic;

KW antiarteriosclerotic; ophthalmological; cytotoxic T cell; CTL; VEGFR-2;

KW vascular endothelial growth factor receptor 2; KDR peptide.

OS Synthetic.

PN WO2004024766-A1.

PD 25-MAR-2004.

PF 12-SEP-2003; 2003WO-JP011722.

PR 12-SEP-2002; 2002JP-00267285.

PR 07-MAR-2003; 2003JP-00062003.

PR 11-JUN-2003; 2003JP-00167042.

PA (ONCO-) ONCOTHERAPY SCI INC.

PA (TAHA/) TAHARA H.

XX Tahara H, Wada S, Tsunoda T;

XX WPI; 2004-295074/27.

XX Nonapeptides comprising cytotoxic T-cell inducing activity, useful as vaccine for treating proliferation of malignant tumors.

XX Example 6; Page 87, 100pp; Japanese.

CC The present invention relates to nona- and decapeptides comprising  
CC cytotoxic T-cell (CTL) inducing activity which are KDR peptides binding  
CC to HLA-A2402, containing partial sequences of receptor protein VEGFR-2  
CC (Vascular endothelial growth factor receptor 2), or derived from these  
CC sequences by addition or substitution of one or more amino acid residues.  
CC The peptides can be used to produce a cancer vaccine for treatment and  
CC prevention of cancer (including stomach, duodenum, colon, lung, breast,  
CC prostate and brain cancer) and suppression of proliferation and  
CC metastasis of malignant tumours. The vaccine is also useful for treatment  
CC of diabetic retinopathy, chronic rheumatoid arthritis and  
CC atherosclerosis. The present sequence is a peptide of the invention.

XX Sequence 10 AA;

Query Match 80.8%; Score 42; DB 8; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.98;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 LLMWFSL 10  
|||  
1 LLMWFSL 8

RESULT 12  
AAG94960  
ID AAG94960 standard; peptide; 10 AA.

XX AAG94960;

XX 18-SEP-2001 (first entry)

XX Human complementary peptide, SEQ ID NO: 1154.

XX Human; complementary peptide; ligand; drug discovery; drug design.

XX Homo sapiens.

XX WO200142277-A2.

XX 14-JUN-2001.

XX 13-DEC-2000; 2000WO-GB004776.

XX 13-DEC-1999; 99GB-00029464.

XX (PROT-) PROTEOM LTD.

XX Roberts GW, Heal JR;

XX WPI, 2001-408419/43.

XX A set of peptide ligands consisting of specific complementary peptides to  
XX proteins encoded by genes of the human genome; useful in an assay for  
XX screening and identifying of one or more novel peptides which are drug  
XX candidates or pro-drugs.

XX Example 4; Page 205; 646pp; English.

XX The invention relates to a set of complementary peptide ligands generated  
XX from the human genome. The complementary peptides interact with their  
XX relevant target proteins encoded in the human genome. They can be used as  
XX reagents in drug discovery and as lead ligands to facilitate drug design  
XX and development. The present sequence is a complementary peptide provided  
XX in the specification

XX Sequence 10 AA;

Query Match 61.5%; Score 32; DB 4; Length 10;  
Best Local Similarity 83.3%; Pred. No. 56;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 5 WEIIFSL 10  
|||  
1 WEIIFSL 6

RESULT 13

AAW97279  
ID AAW97279 standard; peptide; 9 AA.

XX AAW97279;

XX 17-OCT-2003 (revised)

XX 27-AUG-2003 (revised)

XX 20-MAR-2003 (revised)

XX 10-MAY-1999 (first entry)

XX Cytotoxic T-cell epitope from antigen gp85 of EBV.

XX Cytotoxic T-cell epitope; EBV; structural antigen; EBV infection;  
XX infectious mononucleosis; IM; NPC; HD; Burkitt's lymphoma;  
XX post transplantation lymphoproliferative disease; vaccine.

XX Human herpesvirus 4.

XX WO9902550-A1.

XX 21-JAN-1999.

XX 10-JUL-1998; 98WO-AU000531.

XX 10-JUL-1997; 97AU-00007841.

XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.

XX (CSIR) COMMONWEALTH SCI & IND RES ORG.

XX (UNIV) UNIV MELBOURNE.

XX (HALT-) HALT INST MEDICAL RES WALTER & ELIZA.

XX (CSLC-) CSL LTD.

XX (SHER/) SHERITT M A.

XX Burrows SR, Khanna R;

XX WPI; 1999-120774/10.

XX New Epstein-Barr virus antigens - comprise cytotoxic T-cell epitopes,  
XX used to treat or protect against EBV infection, e.g. in nasopharyngeal  
XX carcinoma or Hodgkin's disease.

XX Example 2; Page 28; 73pp; English.

XX AAW97258-90 represent cytotoxic Epstein-Barr virus (EBV) T-cell epitopes  
XX derived from an EBV structural antigen. The epitopes (and nucleic acids  
XX encoding them) can be used for reducing the risk of EBV infection in a  
XX subject. They can be used for reducing the risk of infectious  
XX mononucleosis (IM) or post transplantation lymphoproliferative disease in  
XX a subject. The epitopes can also be used for treating or preventing NPC  
XX or HD or Burkitt's lymphoma. They can also be used in subunit vaccines  
XX with other antigens, e.g. tetanus toxoid, diphtheria toxoid, Bordetella  
XX pertussis antigens, poliovirus antigens, purified protein derivative  
XX (PPD), gp350 protein, helper epitopes or combinations. (Updated on 20-MAR  
XX -2003 to correct PA field.) (Updated on 27-AUG-2003 to correct OS field.)  
XX (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 9 AA;

Query Match 59.6%; Score 31; DB 2; Length 9;  
Best Local Similarity 83.3%; Pred. No. 1.7e+06;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 VLLMEI 7  
|||  
4 VLLMEV 9

```

RESULT 14
AAB19600 standard; peptide; 10 AA.
XX
XX AAB19600;
XX
XX 22-JAN-2001 (first entry)
XX
XX Human CASB616 epitope binding HLA_A0201.
XX
XX CASB616; EPHB2; ERK; EPH3; EPRT3; DRT; HEK5; EPHB2V;
XX receptor protein tyrosine kinase; human; epitope; colon cancer;
XX ovary cancer; tumour; autoimmune disease; vaccine; therapy; diagnosis;
XX HLA-A0201.
XX
XX Homo sapiens.
XX
XX WO200053216-A2.
XX
XX 14-SEP-2000.
XX
XX 26-FEB-2000; 2000WO-EP001587.
XX
XX 05-MAR-1999; 99GB-00005124.
XX
XX (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Vinals y De Bassols;
XX
XX WPI; 2000-587384/55.
XX
XX Vaccine composition for treating ovarian and colon cancer, comprises
XX CASB616 polypeptides, polynucleotides or antigen presenting cells
XX expressing the polypeptides.
XX
XX Example 6; Page 38; 57pp; English.
XX
XX The present sequence is that of a human CASB616 epitope that binds the
XX HLA-A0201 allele. The epitope corresponds to amino acid residues 810-819
XX of the CASB616 polypeptide (see AAB19590 and AAB19591). It is 1 of 10
XX epitopes (see AAB19592) identified during an analysis of human cellular
XX immune responses to CASB616. CASB616 is a tumour specific antigen.
XX CASB616 polypeptides and polynucleotides are used in claimed vaccine
XX compositions for specific prophylactic or therapeutic immunization
XX against tumours, especially colon cancer (claimed) and ovarian cancer, or
XX against autoimmune disease. They can also be used to diagnose the
XX occurrence of tumour cells
XX
XX Sequence 10 AA;
SQ
Query Match 59.6%; Score 31; DB 3; Length 10;
Best Local Similarity 42.9%; Pred. No. 84;
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 GVTLMWEI 7
DB 4 GVTMWEV 10

```

```

RESULT 15
AAY46428
ID AAY46428 standard; peptide; 9 AA.
XX
XX AAY46428;
XX
XX 01-DEC-1999 (first entry)
XX
XX Immunogenic peptide having a human leukocyte antigen binding motif #1039.
XX
XX Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
XX immune response; T cell activation; major histocompatibility complex;
XX cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
XX

```

```

XX prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
XX vaccine; immunisation.
XX
XX Synthetic.
XX
XX Homo sapiens.
XX
XX WO9945954-A1.
XX
XX 16-SEP-1999.
XX
XX 13-MAR-1998; 98WO-US005039.
XX
XX 13-MAR-1998; 98WO-US005039.
XX
XX (EPI-M-) EPIMUNE INC.
XX
XX Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;
XX WPI; 1999-551214/46.
XX
XX New immunogenic peptides with HLA binding motif, useful in treatment and
XX diagnosis of cancers and viral diseases.
XX
XX Claim 1; Page 71; 150pp; English.
XX
XX AAY45390 to AAY48214 represent specifically claimed immunogenic peptides
XX having a human major histocompatibility complex (MHC) Class I (also known
XX as human leukocyte antigen (HLA)) binding motif. The immunogenic peptides
XX can bind to a specific HLA allele (i.e. HLA-A subtypes HLA-A2.1, A1, A3.2
XX or A24.1 or HLA-B or C) and induce a cytotoxic T cell response against
XX the antigen from which the peptide is derived. Cytotoxic T lymphocytes
XX (CTLs) which destroy antigen-bearing cells are normally induced by an
XX antigen in the form of a peptide fragment bound to a HLA molecule, rather
XX than the intact foreign antigen itself, and are particularly important in
XX tumour rejection and in fighting viral infections. The peptides are
XX therefore useful therapeutically to treat or prevent viral infections and
XX cancers in mammals (especially humans) e.g. prostate cancer, hepatitis B
XX and C, AIDS, and renal carcinoma. They can be administered as vaccines to
XX elicit an immune response in individuals susceptible or otherwise at risk
XX of viral infection or cancer, or used to treat chronic or acute
XX conditions. They are also useful diagnostically, and can be used to
XX induce a cytotoxic T cell response, by contacting a cytotoxic T cell with
XX the peptide e.g. to produce CTLs ex vivo for infusion back into a
XX patient. The polynucleotides encoding the immunogenic peptides are also
XX useful therapeutically and for immunisation as above
XX
XX Sequence 9 AA;
SQ
Query Match 53.8%; Score 28; DB 2; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.7e+06;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 GVTLMWEI 7
DB 3 GVTMWEI 9

```

Search completed: November 30, 2004, 08:09:35  
Job time : 58.333 secs

Tue Nov 30 08:49:19 2004

us-10-008-377a-1.open.rpr

BEST AVAILABLE COPY

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 ; Search time 20 Seconds  
(without alignments)  
48.108 Million cell updates/sec

Title: US-10-008-377A-1

Perfect score: 47

Sequence: 1 SLAMDLLHV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1.\*  
2: p1r2.\*  
3: p1r3.\*  
4: p1r4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	100.0	680	4	138491
2	47	100.0	1621	2	T30200
3	37	78.7	311	2	S65091
4	36	76.6	1039	2	F71427
5	36	76.6	4717	2	T41581
6	35	74.5	232	2	H69294
7	35	74.5	441	2	A41591
8	35	74.5	442	1	S13425
9	35	74.5	442	1	JO042
10	35	74.5	442	2	I57950
11	34	72.3	493	2	S63060
12	34	70.2	145	2	T08407
13	33	70.2	290	2	AD5334
14	33	70.2	340	1	OKBYR1
15	33	70.2	391	2	S39816
16	32	68.1	102	2	D71042
17	32	68.1	177	2	G75013
18	32	68.1	190	2	AF1793
19	32	68.1	272	2	AP3515
20	32	68.1	346	2	H70722
21	32	68.1	384	1	SYEC5M
22	32	68.1	384	2	AB0877
23	32	68.1	384	2	B91106
24	32	68.1	384	2	E85951
25	32	68.1	384	2	AC0114
26	32	68.1	409	2	C87319
27	32	68.1	421	2	E95534
28	32	68.1	436	2	G97701
29	32	68.1	454	2	S58148

30	32	68.1	939	2	C70876	hypothetical prote
31	32	68.1	1047	2	S19508	MSH3 protein - Yea
32	31	66.0	83	2	S28120	gas-vesicle operon
33	31	66.0	135	2	E97460	hypothetical prote
34	31	66.0	148	2	S07196	meiosis and sporul
35	31	66.0	158	2	T29432	hypothetical prote
36	31	66.0	233	2	A13495	probable hydrolase
37	31	66.0	246	2	T05961	1,3-beta-glucanase
38	31	66.0	246	2	T05955	hypothetical prote
39	31	66.0	299	2	C72242	hypothetical prote
40	31	66.0	340	2	UC7695	G protein-coupled
41	31	66.0	391	2	A11701	hypothetical prote
42	31	66.0	396	2	H83576	methionine adenosy
43	31	66.0	421	2	C87545	hypothetical prote
44	31	66.0	561	2	G95367	probable methyl-ac
45	31	66.0	592	2	T25837	hypothetical prote

#### ALIGNMENTS

RESULT 1  
138491 nucleophosmin/anaplastic lymphoma kinase mutant fusion protein - human  
C/Species: Homo sapiens (man)  
C/Date: 15-Feb-1996 #sequence\_revision 15-Feb-1996 #text\_change 20-Apr-2000  
C/Accession: 138491  
R/Morris, S.W.; Kirstein, M.N.; Valentine, M.B.; Dittmer, K.G.; Shapiro, D.N.; Saltman,  
Science 263, 1281-1284, 1994  
A/Title: Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's  
A/Reference number: A53182; M01D:94167588; PMID:8122112  
A/Accession: 138491  
A/Molecule type: mRNA  
A/Residues: 1-680 <MOR>  
A/Cross-references: EMBL:U04946; NID:G609341; PIDN:AA58698.1; PID:G609342  
C/Comment: This sequence is the chimeric product of a translocation mutation.  
C/Genetics:  
A/Gene: NPM1/ALK  
A/Map position: 5/2p23-2p23  
C/Keywords: fusion protein

Query Match 100.0%; Score 47; DB 4; Length 680;  
Best Local Similarity 100.0%; Pred. No. 0.13;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLLHV 10  
Db 280 SLAMDLLHV 289

#### RESULT 2

T30200 protein-tyrosine kinase (EC 2.7.1.112) alk - mouse  
N/Alternate names: anaplastic lymphoma kinase  
C/Species: Mus musculus (house mouse)  
C/Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T30200  
R/Iwanaga, T.; Fujimoto, J.; Wen, D.; Cupples, R.; Bucay, N.; Arakawa, T.; Mori, S.; Ra  
Oncogene 14, 439-449, 1997  
A/Title: Molecular characterization of ALK, a receptor tyrosine kinase expressed specif  
A/Reference number: Z20774; M01D:97178853; PMID:9053841  
A/Accession: T30200  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-1621 <1MA>  
A/Cross-references: UNIPROT:P97793; EMBL:DB3002; NID:G1864006; PIDN:BA11673.1; PID:G18  
A/Experimental source: brain and testis  
C/Genetics:  
A/Gene: alk  
C/Function:  
A/Description: may play an important role in development of the brain  
C/Keywords: ATP, phosphotransferase; tyrosine-specific protein kinase

Query Match 100.0%; Score 47; DB 2; Length 1621;  
 Best Local Similarity 100.0%; Pred. No. 0.34;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
 DB 1224 SLAMDLHLV 1233

## RESULT 3

S65091

geranylgeranyltransferase type II beta chain - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe

C>Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004

C:Accession: S65091; T37741

C:Godfrey, R.; Davey, J.

Yeast 12, 479-483, 1996

A:Title: Sequence of pGbl, a gene for the beta subunit of the type-II geranylgeranyltran

A:Reference number: S65091; MUID:96310627; PMID:8740421

A:Status: preliminary

A:Accession: S65091

A:Molecule type: DNA

A:Residues: 1-311 <GOD>

A:Cross-references: UNIPROT:P46960; EMBL:X82183; NID:G1033065; PIDN:CAA63094.1; PID:G103

R; Rieger, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, January 1999

A:Reference number: 221743

A:Accession: T37741

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-311 <RIF>

A:Cross-references: EMBL:AL035248; PIDN:CAA22847.1; GSPDB:GN00066; SPDB:SPAC167.02

A:Experimental source: strain 972h; cosmid c167

C:Genetics:

A:Gene: SPAC167.02

A:Map position: 1

A:Introns: 18/3; 69/1; 108/1; 180/1; 224/3

C:Superfamily: DEPI protein

QY 2 LAMLDLHV 10  
 DB 91 LAMLDLHV 99

Query Match 78.7%; Score 37; DB 2; Length 311;  
 Best Local Similarity 88.9%; Pred. No. 6.7;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 91 LAMLDLHV 99

Query Match 78.7%; Score 37; DB 2; Length 311;  
 Best Local Similarity 88.9%; Pred. No. 6.7;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

hypothetical protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

A:Variety: Columbia

C>Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 16-Aug-2004

C:Accession: F71427

R; Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk

P.; Medler, H.; Weiler, E.; Wambutt, R.; Weitzel, T.; Pohl, T.M.; Terry, N.; Giel

avannah, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaefer, M.; Funk, B.

Nature 391, 485-488, 1998

A:Authors: Weiler, E.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk

avannah, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaefer, M.; Funk, B.

A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal

A:Reference number: A74400; MUID:9812113; PMID:9461215

A:Accession: F71427

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1039 <BEV>

A:Cross-references: UNIPROT:O23461; GB:Z97340; NID:G2244950; PID:G2244971

C:Genetics:

A:Map position: 4COP9-4G3845

C:Superfamily: Galactokinase with glycosyltransferase N-terminal domain

Query Match 76.6%; Score 36; DB 2; Length 1039;  
 Best Local Similarity 77.8%; Pred. No. 3.9;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLHV 9  
 DB 514 SLAMDLHV 522

## RESULT 5

T41581

hypothetical coiled-coil protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C:Accession: T41581

R; Murphy, L.; Harris, D.; Wood, V.; Lyne, M.H.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, September 1998

A:Reference number: Z22002

A:Accession: T41581

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-4717 <MUR>

A:Cross-references: UNIPROT:Q94248; EMBL:AL031546; PIDN:CAA20864.1; GSPDB:GN00068; SPDB

A:Experimental source: strain 972h(-)

C:Genetics:

A:Gene: SPDB:SPCC737.08

A:Map position: 3

C:Superfamily: mda5in (AAA Arpase with von Willebrand factor type A (vwa) domain)

QY 2 LAMLDLHV 10  
 DB 2516 LAMLDLHV 2524

Query Match 76.6%; Score 36; DB 2; Length 4717;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
 DB 2516 LAMLDLHV 2524

Query Match 76.6%; Score 36; DB 2; Length 4717;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

conserved hypothetical protein AF0360 - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004

C:Accession: H69294

R; Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirschner, E.F.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Overbeek, R.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes, S

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae

A:Reference number: A69250; MUID:98049343; PMID:9389475

A:Accession: H69294

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-232 <KLE>

A:Cross-references: UNIPROT:Q29887; GB:AB001079; GB:AB000782; NID:G2689402; PIDN:AA8908

Query Match 74.5%; Score 35; DB 2; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AMMDLHV 9  
 DB 109 AMMDLHV 115

Query Match 74.5%; Score 35; DB 2; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AMMDLHV 9  
 DB 109 AMMDLHV 115

Query Match 74.5%; Score 35; DB 2; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AMMDLHV 9  
 DB 109 AMMDLHV 115

Query Match 74.5%; Score 35; DB 2; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AMMDLHV 9  
 DB 109 AMMDLHV 115

Query Match 74.5%; Score 35; DB 2; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AMMDLHV 9  
 DB 109 AMMDLHV 115

Query Match 74.5%; Score 35; DB 2; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AMMDLHV 9  
 DB 109 AMMDLHV 115

Query Match 74.5%; Score 35; DB 2; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AMMDLHV 9  
 DB 109 AMMDLHV 115

Query Match 74.5%; Score 35; DB 2; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AMMDLHV 9  
 DB 109 AMMDLHV 115

Query Match 74.5%; Score 35; DB 2; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

R,Saito, Y.; Mizuno, T.; Itakura, M.; Suzuki, Y.; Ito, T.; Hagiwara, H.; Hirose, S.  
 J. Biol. Chem. 266, 23433-23437, 1991  
 A>Title: Primary structure of bovine endothelin ET-B receptor and identification of sign  
 A/Reference number: A41591; MUID:92078223; PMID:1660473  
 A/Accession: A41591  
 A/Molecule type: mRNA  
 A/Residues: 1-441 <SAI>  
 A/Cross-references: UNIPROT:P28088; GB:D90456; NID:g217553; PIDN:BA14422.1; PID:g217554  
 R,Hagiwara, H.; Kozuka, M.; Sakaguchi, H.; Eguchi, S.; Ito, T.; Hirose, S.  
 J. Cardiovasc. Pharmacol. 17(Suppl.7), S117-S118, 1991  
 A>Title: Separation and purification of 34- and 52-kDa species of bovine lung endothelin  
 A/Reference number: A61251; MUID:92219666; PMID:1725302  
 A/Accession: A61251  
 A/Molecule type: protein  
 A/Residues: 304-315/424-432 <HAG>  
 R,Kozuka, M.; Ito, T.; Hirose, S.; Lodhi, K.M.; Hagiwara, H.  
 J. Biol. Chem. 266, 16892-16896, 1991  
 A>Title: Purification and characterization of bovine lung endothelin receptor.  
 A/Reference number: A41184; MUID:91358493; PMID:1653249  
 A/Accession: A41184  
 A/Molecule type: protein  
 A/Residues: 144-127; S, 263-269; 304-316/418-421; 424-432 <KOZ>  
 R,Hick, S.; Heidemann, T.; Soskic, V.; Mueller-Esterl, W.; Godovac-Zimmermann, J.  
 Eur. J. Biochem. 234, 251-257, 1995  
 A>Title: Isolation of the endothelin B receptor from bovine lung. Structure, signal sequ  
 A/Reference number: S63513; MUID:96096746; PMID:8529649  
 A/Accession: S63513  
 A/Molecule type: protein  
 A/Residues: 1-'X',3-7,'X',9-10;25-35;36-42;73-78;79-87;284-290;291-296,'X',298 <HIC>  
 C/Superfamily: endothelin receptor B  
 C/Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein  
 F/1-26/Domain: signal sequence #status experimental <SIG>  
 F/27-441/Product: endothelin receptor B #status experimental <MAT>  
 F/101-126/Domain: transmembrane #status predicted <TM1>  
 F/117-162/Domain: transmembrane #status predicted <TM2>  
 F/175-196/Domain: transmembrane #status predicted <TM3>  
 F/218-442/Domain: transmembrane #status predicted <TM4>  
 F/271-295/Domain: transmembrane #status predicted <TM5>  
 F/324-449/Domain: transmembrane #status predicted <TM6>  
 F/362-389/Domain: transmembrane #status predicted <TM7>  
 F/60/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 74.5%; Score 35; DB 2; Length 441;  
 Best Local Similarity 70.0%; Pred. No. 25;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Cy 1 SLAMDLHLV 10  
 Db 141 SLALGDLHLI 150

RESULT 8  
 S13425  
 endothelin receptor ETB - rat  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 02-Dec-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
 C/Accession: S13425; A49197  
 R,Sakurai, T.; Yanagisawa, M.; Takawa, Y.; Miyazaki, H.; Kimura, S.; Goto, K.; Maekaki, T  
 Nature 348, 732-735, 1990  
 A>Title: Cloning of a cDNA encoding a non-isopeptide-selective subtype of the endothelin  
 A/Reference number: S13425; MUID:91080924; PMID:2175397  
 A/Accession: S13425  
 A/Molecule type: mRNA  
 A/Residues: 1-441 <SAK>  
 A/Cross-references: UNIPROT:P21451; GB:X57764; NID:G56122; PIDN:CAA40916.1; PID:G56123  
 R,Hori, S.; Komatsu, Y.; Shigemoto, R.; Mizuno, N.; Nakanishi, S.  
 Endocrinology 130, 1885-1895, 1992  
 A>Title: Distinct tissue distribution and cellular localization of two messenger ribonuc  
 A/Reference number: A49197; MUID:92191882; PMID:1312429  
 A/Accession: A49197  
 A/Status: preliminary; not compared with conceptual translation  
 A/Molecule type: nucleic acid  
 A/Residues: 1-65; 'SSAP', 69-441 <HOR>

A/Experimental source: lung  
 A/Note: sequence extracted from NCBI backbone (NCBI:P:89468)  
 C/Superfamily: endothelin receptor B  
 C/Keywords: G protein-coupled receptor; transmembrane protein

Query Match 74.5%; Score 35; DB 2; Length 441;  
 Best Local Similarity 70.0%; Pred. No. 25;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Cy 1 SLAMDLHLV 10  
 Db 141 SLALGDLHLI 150

RESULT 9  
 JQ1042  
 endothelin receptor B precursor - human  
 C/Species: Homo sapiens (man)  
 C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
 C/Accession: A46609; JQ1042; JN0072; JQ1082; B45481  
 R,Arai, H.; Nakao, K.; Takaya, K.; Hosoda, K.; Ogawa, Y.; Nakanishi, S.; Imura, H.  
 J. Biol. Chem. 268, 3463-3470, 1993  
 A>Title: The human endothelin-B receptor gene. Structural organization and chromosomal  
 A/Reference number: A46609; MUID:91155196; PMID:8429023  
 A/Accession: A46609  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-442 <ARA>  
 A/Cross-references: UNIPROT:P24530; GB:D13168; NID:G285924; PIDN:BA402445.1; PID:G285924  
 R,Ogawa, Y.; Nakao, K.; Arai, H.; Nakagawa, O.; Hosoda, K.; Suga, S.; Nakanishi, S.; Im  
 Biochem. Biophys. Res. Commun. 178, 248-255, 1991  
 A>Title: Molecular cloning of a non-isopeptide-selective human endothelin receptor.  
 A/Reference number: JQ1042; MUID:91298956; PMID:1648908  
 A/Accession: JQ1042  
 A/Molecule type: mRNA  
 A/Residues: 1-442 <OGA>  
 A/Cross-references: GB:D90402; NID:G219651; PIDN:BA14398.1; PID:G219652  
 A/Experimental source: placenta  
 R,Nakamura, M.; Takayanagi, R.; Sakai, Y.; Sakamoto, S.; Hagiwara, H.; Mizuno, T.; Saito  
 Biochem. Biophys. Res. Commun. 177, 34-39, 1991  
 A>Title: Cloning and sequence analysis of a cDNA encoding human non-selective type of en  
 A/Reference number: JN0072; MUID:91254298; PMID:1710450  
 A/Accession: JN0072  
 A/Molecule type: mRNA  
 A/Residues: 1-442 <NKA>  
 A/Cross-references: GB:M74921; NID:G182275; PIDN:AAA58465.1; PID:G182276  
 A/Experimental source: liver  
 R,Sakamoto, A.; Yanagisawa, M.; Sakurai, T.; Takawa, Y.; Yanagisawa, H.; Masaki, T.  
 Biochem. Biophys. Res. Commun. 178, 656-663, 1991  
 A>Title: Cloning and functional expression of human cDNA for the ETB endothelin recepto  
 A/Reference number: JQ1082; MUID:91315496; PMID:1713452  
 A/Accession: JQ1082  
 A/Molecule type: mRNA  
 A/Residues: 1-9,'P',11-442 <SAK>  
 A/Cross-references: GB:S44866; NID:G233233; PIDN:AA19411.1; PID:G233234  
 A/Experimental source: jejunum  
 R,Elshourbagy, N.A.; Korman, D.R.; Wu, H.L.; Sylvester, D.R.; Lee, J.A.; Nuthalaganti, B  
 J. Biol. Chem. 268, 3873-3879, 1993  
 A>Title: Molecular characterization and regulation of the human endothelin receptors.  
 A/Reference number: A45481; MUID:93179382; PMID:8440662  
 A/Accession: B45481  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-442 <ELB>  
 A/Cross-references: GB:L06623; NID:G181958; PIDN:AAA52342.1; PID:G181959  
 A/Experimental source: lung  
 A/Note: sequence extracted from NCBI backbone (NCBI:N:125754, NCBI:P:125755)  
 C/Comment: This endothelin receptor is non-isopeptide selective.  
 C/Genetics:  
 A/Gene: GDB:EDNRB; ETB  
 A/Cross-references: GDB:129075; OMIM:131244  
 A/Map position: 13q22-13q22  
 C/Superfamily: endothelin receptor B

C/Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane protein  
 F1-26/Domain: signal sequence #status predicted <SIG>  
 F127-442/Product: endothelin receptor B #status predicted <MAT>  
 F102-127/Domain: transmembrane #status predicted <TM1>  
 F1139-153/Domain: transmembrane #status predicted <TM2>  
 F1176-197/Domain: transmembrane #status predicted <TM3>  
 F1219-243/Domain: transmembrane #status predicted <TM4>  
 F1272-236/Domain: transmembrane #status predicted <TM5>  
 F1325-350/Domain: transmembrane #status predicted <TM6>  
 F1363-389/Domain: transmembrane #status predicted <TM7>  
 F159/Binding site: carbohydrate (asn) (covalent) #status predicted  
 F1305,413,419,435,436,440/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 74.5% Score 35; DB 1; Length 442;  
 Best Local Similarity 70.0%; Pred. No. 25;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLIHV 10  
 |||:||||:  
 Db 142 SLALGDLHI 151

## RESULT 10

157950 nonselective-type endothelin receptor - rat  
 C/Species: Rattus sp. (rat)  
 C/Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 16-Jul-1999

C/Accession: 157950  
 R/Cheng, H.F.; Su, Y.M.; Yeh, J.R.; Chang, K.J.  
 Mol. Pharmacol. 44, 533-538, 1993

A/Title: Alternative transcript of the nonselective-type endothelin receptor from rat brain  
 A/Reference number: 157950; MUID:93382424; PMID:8371713

A/Accession: 157950

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-442 <RES>

A/Cross-references: GB:S65355; NID:G410693; PIDN:AB28172.1; PID:G410693

C/Superfamily: endothelin receptor B

Query Match 74.5% Score 35; DB 2; Length 442;  
 Best Local Similarity 70.0%; Pred. No. 25;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLIHV 10  
 |||:||||:  
 Db 142 SLALGDLHI 151

## RESULT 11

S63060 hypothetical protein YNL119w - yeast (Saccharomyces cerevisiae)

N/Alternate names: hypothetical protein NI913

C/Species: Saccharomyces cerevisiae

C/Date: 27-Apr-1996 #sequence\_revision 03-May-1996 #text\_change 09-Jul-2004

C/Accession: S63060; S67339

R/De Antoni, A.; D'Angelo, M.; Dal Pero, F.; Sartorello, F.; Pandolfo, D.; Lanfranchi, G.

submitted to the Protein Sequence Database, April 1996

A/Reference number: S63047

A/Accession: S63060

A/Molecule type: DNA

A/Residues: 1-493 <DEA>

A/Cross-references: UNIPROT:P53923; EMBL:Z71396; NID:G1302046; PID:E239801; PID:G1302048

A/Experimental source: strain S288C

R/d'Antoni, A.; d'Angelo, M.; dal Pero, F.; Sartorello, F.; Pandolfo, D.; Lanfranchi, G.

submitted to the EMBL Data Library, February 1996

A/Description: The DNA sequence of cosmid 14-13b from chromosome XIV of Saccharomyces cerevisiae

A/Reference number: S67327

A/Accession: S67339

A/Molecule type: DNA

A/Residues: 1-493 <DNA>

A/Cross-references: EMBL:Z69382; NID:G1183941; PID:E221827; PID:G1183954

C/Genetics: SGD:S0005063

A/Map position: 14L  
 A/Note: YNL119w

Query Match 72.3% Score 34; DB 2; Length 493;  
 Best Local Similarity 66.7%; Pred. No. 45;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLIHV 9  
 |||:||||:  
 Db 76 SLVMDLIH 84

## RESULT 12

T08407

clathrin coat assembly protein homolog F18B3.140 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004

C/Accession: T08407

R/Quetier, F.; Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Salanoubat

submitted to the Protein Sequence Database, May 1999

A/Reference number: Z16409

A/Accession: T08407

A/Molecule type: DNA

A/Residues: 1-145 <QUB>

A/Cross-references: UNIPROT:Q9SVL3; EMBL:AL049862; GSPDB:GN00061; ATSP:F18B3.140

A/Experimental source: cultivar Columbia; BAC clone F18B3

C/Genetics:

A/Map position: 3

A/Introns: 23/3; 38/1; 58/3; 91/1; 114/3

C/Superfamily: mouse clathrin-associated protein 19

Query Match 70.2% Score 33; DB 2; Length 145;  
 Best Local Similarity 77.8%; Pred. No. 20;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMDLIHV 10  
 |||:||||:  
 Db 83 LAMDLIHV 91

## RESULT 13

AD3534 antibiotic-induced protein drp35 [imported] - Brucella melitensis (strain 16M)

C/Species: Brucella melitensis

C/Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 09-Jul-2004

C/Accession: AD3534

R/DeIvecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A/Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A/Reference number: AD3252; PMID:11756688

A/Accession: AD3534

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-230 <KUR>

A/Cross-references: UNIPROT:Q8YDH8; GB:AB008918; PIDN:AAL53439.1; PID:G17984337; GSPDB:C

A/Experimental source: strain 16M

C/Genetics:

A/Map position: 11

Query Match 70.2% Score 33; DB 2; Length 290;  
 Best Local Similarity 66.7%; Pred. No. 42;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLIHV 9  
 |||:||||:  
 Db 26 SLAVDLIH 34

## RESULT 14

OKBYR1

protein kinase byr1 (EC 2.7.1.1) - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004

C:Accession: S00473; T38054

R:Naclin-Davis, S.A.; Naclim, A.

EMBO J. 7, 985-993, 1988

A:Title: A gene which encodes a predicted protein kinase can restore some functions of

A:Reference number: S00473; PMID:8826442; PMID:3042386

A:Accession: S00473

A:Molecule type: DNA

A:Residues: 1-340 <NAD>

A:Cross-references: UNIPROT:P10506; EMBL:X07445; NID:G4918; PIDN:CAA30326.1; PID:G4919

R:Lyse, G.; Churcher, C.M.; Barrall, B.G.; Rajadream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, February 1995

A:Reference number: 221765

A:Accession: T38054

A:Status: preliminary; translated from GH/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-340 <LYE>

A:Cross-references: EMBL:Z69239; PIDN:CAA9322.1; GSPDB:GN00066; SPDB:SPAC1D4.13

A:Experimental source: strain 972-; cosmid c1D4

C:Comment: Inactivation of this gene does not interfere with mitotic growth but prevents

C:Genetics:

A:Gene: byr1

A:Map position: 1

C:Superfamily: kinase-related transforming protein; protein kinase homology

C:Keywords: ATP; cell cycle control; phosphoprotein; phosphotransferase; serine/threonine

F:64.320/Domain: protein kinase homology <KIN>

F:72-80/Region: protein kinase ATP-binding motif

F:93/Active site: Lys #status predicted

Query Match 70.2%; Score 33; DB 1; Length 340;

Best Local Similarity 66.7%; Pred. No. 49;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLIH 9

Db 267 SIGILDLIH 275

RESULT 15

S39816

lysine acetyltransferase - Yeast (Yarrowia lipolytica)

C:Species: Yarrowia lipolytica; Candida lipolytica

C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004

C:Accession: S39816; S18927

R:Beckerich, J.M.; Lambert, M.; Gallardin, C.

Curr. Genet. 25, 24-29, 1994

A:Title: LYC1 is the structural gene for lysine N-6-acetyl transferase in yeast.

A:Reference number: S39816; PMID:9436371; PMID:8082161

A:Accession: S39816

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-391 <BEC>

A:Cross-references: UNIPROT:P41929; EMBL:X63548; NID:G5246; PIDN:CAA45112.1; PID:G5247

C:Genetics:

A:Gene: lyc1

C:Superfamily: Yarrowia lipolytica lysine acetyltransferase

Query Match 70.2%; Score 33; DB 2; Length 391;

Best Local Similarity 66.7%; Pred. No. 57;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLIH 9

Db 166 SLRLMDLIH 174

Search completed: November 30, 2004, 07:41:22

Job time : 21 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 ; Search time 76.4103 Seconds  
(without adjustment)

75.301 Million cell updates/sec

Title:	US-10-008-377A-1
Perfect score:	47
Sequence:	1 SLAMDDLHV 10

Scoring table: BLOSUM62

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

```
Database : uniProt_02:*
1: uniProt_sprot:*
2: uniProt_trembl:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	47	100.0	351	2	Q6CJH1	Q6CJH1 rattus norv
2	47	100.0	803	2	Q6RDS5	Q6RDS5 homo sapien
3	47	100.0	1820	1	AAK_HUMAN	Q9UMW3 homo sapien
4	47	100.0	1621	1	AAK_MOUSE	P97793 mus musculu
5	37	78.7	311	1	PTB1_SCHPO	P46960 schistosac
6	37	78.7	792	1	Q6EC13	Q6EC13 shewanella
7	37	78.7	912	2	Q6JFV0	Q6JFV0 brachydanio
8	36	76.6	376	2	Q5HX1	Q5HX1 cyzsa satili
9	36	76.6	461	2	Q7SBK0	Q7SBK0 ashyba goss
10	36	76.6	461	2	AA551497	AA551497 ashyba g
11	36	76.6	763	2	Q7SAB6	Q7SAB6 neurospora
12	36	76.6	872	2	Q6RUM3	Q6RUM3 xanthomonas
13	36	76.6	989	2	Q6J723	Q6J723 arabidopsis1
14	36	76.6	1039	2	Q23461	Q23461 arabidopsis1
15	36	76.6	1066	2	Q7QRY1	Q7QRY1 giardia lam
16	36	76.6	4717	2	Q94248	Q94248 schistosac
17	35	74.5	152	2	Q9FIS6	Q9FIS6 streptococc
18	35	74.5	208	2	Q9FIS3	Q9FIS3 streptococc
19	35	74.5	232	2	Q29887	Q29887 archaeoglob
20	35	74.5	276	2	Q9E1T3	Q9E1T3 streptococc
21	35	74.5	276	2	Q23175	Q23175 streptococc
22	35	74.5	312	2	Q6BUE1	Q6BUE1 mus musculu
23	35	74.5	347	1	ETBR_COTJA	Q90338 coturnix co
24	35	74.5	356	2	Q6JHY4	Q6JHY4 gallus galli
25	35	74.5	382	2	Q7NRP3	Q7NRP3 chromobact
26	35	74.5	426	1	ETBR_CANFA	P56497 canis fami
27	35	74.5	426	2	Q6MX61	Q6MX61 azoarcus sp
28	35	74.5	426	2	Q9DGM2	Q9DGM2 brachydanio
29	35	74.5	426	2	CAF21977	CAF21977 azoarcus
30	35	74.5	434	2	Q6MTZ5	Q6MTZ5 canis fami
31	35	74.5	441	1	ETBR_BOVIN	P28088 bos taurus

32	35	74.5	44.1	2	Q9N0W7
33	35	74.5	44.1	1	ETBR_HUMAN
34	35	74.5	44.2	1	ETBR_MOUSE
35	35	74.5	44.2	1	ETBR_PAT
36	35	74.5	44.2	2	Q9UD23
37	35	74.5	44.2	2	Q8KIV9
38	35	74.5	44.2	2	AAS31516
39	35	74.5	44.2	2	BAC36337
40	35	74.5	44.2	2	BAC38409
41	35	74.5	44.2	2	BAC38908
42	35	74.5	44.2	2	BAC39465
43	35	74.5	44.3	1	ETBR_HORSE
44	35	74.5	44.3	1	ETBR_PIG
45	35	74.5	61.2	2	Q8VXF6
					Q8VXF6 lupinus lute

## ALIGNMENTS

```

RESULT 1
OCSTH1
ID OCSTH1 PRELIMINARY; PRT; 351 AA.
AC OCSTH1;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Anaplastic lymphoma kinase (Fragment).
GN NameAalk;
OS Rattus norvegicus (Rat).
CC Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RA Aashita K., Sato H., Yamasaki C., Tateno C., Katoka M., Shikawa M.,
RA Katayama S., Yoshizato K.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB073169; BAC21663.1; -
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0004686; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase-like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002011; ReceptTykKinII.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008265; Tyr_kinase_AS.
DR Pfam; PF00069; Pkinase; 1
DR PRINTS; PRO0109; TYRKINASE.
DR ProDom; PD000001; Prot_Kinase; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN II; UNKNOWN_1.
KW ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.
FT NON_TER 1
FT TER 351
SQ SEQUENCE 351 AA; 38984 MW; 6ADCEC220C76B125 CRC64;
Query Match 100.0%; Score 47; DB 2; Length 351;
Best Local Similarity 100.0%; Prod. No. 0.41;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

AC Q8TUD5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 26, Last annotation update)
DE TRK-fused gene/anaplastic large cell lymphoma kinase extra long
DE form.
GN Name=TRK/ALK fusion;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21940189; PubMed=1194373;
RA Hernandez L., Bea S., Bellonillo B., Pinyol M., Falini B., Ott G.,
RA Pulford K., Rosenwald A., Morris S.W., Fernandez A., Santos E.,
RA Campo E.;
RT "Diversity of genomic breakpoints in TRK-ALK translocations in
RT anaplastic large cell lymphomas: identification of a new TRK-ALK(XL)
RT Am. J. Pathol. 160:1487-1494(2002)."
RL EMBL: AF390893; AAM17922.1; -
DR HSSP: O62838; 1LUF.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO: GO:0018740; F:transferase activity; IEA.
DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR011009; Kinase like.
DR InterPro: IPR000270; CRR_PBL.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002011; ReceptTyknsInt.
DR InterPro: IPR001245; Tyr_kinase.
DR InterPro: IPR008266; Tyr_kinase_AS.
DR Pfam: PF00564; PBL; 1.
DR Pfam: PF00069; Pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00219; Tyrc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKOWN_1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00239; RECEPTOR_TYR_KIN_11; UNKOWN_1.
DR ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.
KW SEQUENCE 803 AA; 88670 MW; 890826D49EB6E9FF CRC64;
SQ
Query Match 100.0%; Score 47; DB 2; Length 803;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLAMF1LHV 10
DB 403 SLAMF1LHV 412

```

```

RA Morris S.W., Naeve C.W., Mathew P., James P.L., Kirstein M.N., Cui X.,
RA Witte D.P.;
RT "ALK, the chromosome 2 gene locus altered by the t(2;5) in non-
RT Hodgkin's lymphoma, encodes a novel neural receptor tyrosine kinase
RT that is highly related to leukocyte tyrosine kinase (LTK).";
RL Oncogene 14:2175-2186(1997).
RN [2]
RP ERRATUM.
RA Morris S.W., Naeve C.W., Mathew P., James P.L., Kirstein M.N., Cui X.,
RA Witte D.P.;
RL Oncogene 15:2883-2883(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97178863; PubMed=9053841;
RA Iwahara T., Fujimoto J., Wen D., Cupples R., Bucay N., Arakawa T.,
RA Mori S., Ratzkin B., Yamamoto T.;
RT "Molecular characterization of ALK, a receptor tyrosine kinase
RT expressed specifically in the nervous system.";
RL Oncogene 14:439-449(1997).
RN [4]
RP PARTIAL SEQUENCE FROM N.A., AND CHROMOSOMAL TRANSLOCATION.
RX MEDLINE=94167588; PubMed=8122112;
RA Morris S.W., Kirstein M.N., Valentine M.B., Dittmer K.G.,
RA Shapiro D.N., Saltman D.L., Look A.T.;
RT "Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in
RT non-Hodgkin's lymphoma.";
RL Science 263:1281-1284(1994).
CC -1- FUNCTION: Orphan receptor with a tyrosine-protein kinase activity.
CC Appears to play an important role in the normal development and
CC function of the nervous system.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed in brain and CNS. Also expressed in
CC the small intestine and testis, but not in normal lymphoid cells.
CC -1- PFM: N-glycosylated.
CC -1- DISEASE: A form of non-Hodgkin's lymphoma is characterized by a
CC chromosomal translocation t(2;5) (p23;q35) that involves NPM1 and
CC ALK.
CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin
CC receptor subfamily.
CC -1- SIMILARITY: Contains 1 LDL-receptor class A domain.
CC -1- SIMILARITY: Contains 2 WAM domains.
CC -1- DATAABAS: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW="http://www.infobio.gen.fr/services/chromocancer/genes/ALK.html".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC
DR EMBL: U62540; AAB71619.1; -.
DR EMBL: U66559; AAC51104.1; -.
DR HSSP: O62838; 1LUF.
DR Genew; HGNC:427; ALK.
DR MIM; 105590; -.
DR GO: GO:0016021; C:integral to membrane; NAS.
DR GO: GO:0007399; P:neurogenesis; NAS.
DR InterPro: IPR011009; Kinase like.
DR InterPro: IPR002172; LDL_receptor_A.
DR InterPro: IPR000998; WAM.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002011; ReceptTyknsInt.
DR InterPro: IPR001245; Tyr_kinase.
DR InterPro: IPR008266; Tyr_kinase_AS.
DR Pfam: PF00069; Pkinase; 1.
DR Pfam: PF00564; PBL; 1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00192; LDLa; 1.

```

DR SMART; SM00219; TykC; 1.  
 DR PROSITE; PS01209; LDLR\_1; FALSE NEG.  
 DR PROSITE; PS00668; LDLR\_2; FALSE NEG.  
 DR PROSITE; PS00740; MAM\_1; FALSE NEG.  
 DR PROSITE; PS00660; MAM\_2; 2.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00101; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE TYR; 1.  
 DR PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
 DR ATP-binding; Chromosomal translocation; Phosphorylation;  
 KW Pto-oncogene; Receptor; Repeat; Signal; Transferrase; Transmembrane;  
 KW Tyrosine-protein kinase.  
 FT SIGNAL 1 18  
 FT CHAIN 19 1620 Potential.  
 FT DOMAIN 19 1038 ALK tyrosine kinase receptor.  
 FT TRANSMEM 1039 1059 Extracellular (Potential).  
 FT DOMAIN 1060 1620 Potential.  
 FT DOMAIN 264 427 Cytoplasmic (Potential).  
 FT DOMAIN 437 473 MAM 1.  
 FT DOMAIN 478 636 LDL-receptor class A.  
 FT DOMAIN 1116 1392 MAM 2.  
 FT NP\_BIND 816 940 Protein kinase.  
 FT BINDING 1122 1130 Gly-rich.  
 FT ACT\_SITE 1150 1150 ATP (By similarity).  
 FT MOD\_RES 1249 1249 ATP (By similarity).  
 FT MOD\_RES 1282 1282 Proton acceptor (By similarity).  
 FT CARBOHYD 169 169 Phosphotyrosine (by autocatalysis) (By  
 FT CARBOHYD 244 244 similarity).  
 FT CARBOHYD 285 285 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 324 324 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 411 411 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 424 424 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 445 445 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 563 563 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 571 571 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 627 627 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 709 709 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 808 808 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 863 863 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 864 864 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 886 886 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 986 986 N-linked (GlcNAc...) (Potential).  
 FT CONFLICT 36 36 S -> P (in Ref. 3).  
 FT CONFLICT 1491 1491 K -> R (in Ref. 3).  
 FT CONFLICT 1529 1529 D -> E (in Ref. 3).  
 SQ SEQUENCE 1620 AA; 176417 MW; A62604B242961B1B CRC64;

Query Match 100.0%; Score 47; DB 1; Length 1620;  
 Best Local Similarity 100.0%; Pred. No. 2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLAMF1LHV 10  
 DB 1220 SLAMF1LHV 1229

RESULT 4  
 ID ALK\_MOUSE STANDARD; PRT; 1621 AA.  
 AC P97793;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUN-2004 (Rel. 44, Last annotation update)  
 DE ALK tyrosine kinase receptor precursor (BC 2.7.1.112) (Anaplastic  
 DE lymphoma kinase).  
 GN Name=Alk;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE=Brain, and Testis;  
 RX MEDLINE=9717863; PubMed=9053841;  
 RA Iwihara T., Fujimoto J., Wen D., Cupples R., Bucay N., Arakawa T.,  
 RA Mori S., Ratzkin B., Yamamoto T.;  
 RT "Molecular characterization of ALK, a receptor tyrosine kinase  
 RT expressed specifically in the nervous system";  
 RL Oncogene 14:439-449(1997).  
 CC -1 FUNCTION: Orphan receptor with a tyrosine-protein kinase activity.  
 CC Appears to play an important role in the normal development and  
 CC function of the nervous system.  
 CC -1 CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1 SUBCELLULAR LOCATION: Type I membrane protein (Probable).  
 CC -1 SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin  
 CC receptor subfamily.  
 CC -1 SIMILARITY: Contains 2 MAM domains.  
 CC -1 SIMILARITY: Contains 2 MAM domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; D83002; BAA1673.1; -  
 DR PIR; T30200; T30200.  
 DR HSRP; P08069; 1JQH.  
 DR MGD; MGI:103305; Alk.  
 DR GO; GO:0005515; Fiprotein binding; IPI.  
 DR InterPro; IPR011009; Kinase\_like.  
 DR InterPro; IPR002172; LDL\_receptor\_A.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR002011; ReceptTyKinsII.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR InterPro; IPR008266; Tyr\_kinase\_AS.  
 DR Pfam; PF00629; MAM\_1.  
 DR Pfam; PF00629; PKinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR PRODOM; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00192; LDla; 1.  
 DR SMART; SM00219; TykC; 1.  
 DR PROSITE; PS01209; LDLR\_1; FALSE NEG.  
 DR PROSITE; PS00668; LDLR\_2; FALSE NEG.  
 DR PROSITE; PS00740; MAM\_1; FALSE NEG.  
 DR PROSITE; PS00660; MAM\_2; 2.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00101; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE TYR; 1.  
 DR PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
 KW ATP-binding; Glycoprotein; Phosphorylation; Receptor; Repeat; Signal;  
 KW Transferrase; Transmembrane; Tyrosine-protein kinase.  
 FT SIGNAL 1 18 Potential.  
 FT CHAIN 19 1621 ALK tyrosine kinase receptor.  
 FT DOMAIN 19 1042 Extracellular (Potential).  
 FT TRANSMEM 1043 1063 Potential.  
 FT DOMAIN 1064 1621 Cytoplasmic (Potential).  
 FT DOMAIN 268 431 MAM 1.  
 FT DOMAIN 441 477 LDL-receptor class A.  
 FT DOMAIN 482 640 MAM 2.  
 FT DOMAIN 820 944 Protein kinase.  
 FT NP\_BIND 1126 1134 Gly-rich.  
 FT BINDING 1154 1154 ATP (By similarity).  
 FT ACT\_SITE 1253 1253 Proton acceptor (By similarity).  
 FT MOD\_RES 1286 1286 Phosphotyrosine (by autocatalysis) (By  
 FT CARBOHYD 174 174 similarity).  
 FT CARBOHYD 248 248 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 289 289 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 328 328 N-linked (GlcNAc...) (Potential).

FT CAROCHD 415 415 N-linked (GlcNAc. . .) (Potential).  
 FT CAROCHD 428 428 N-linked (GlcNAc. . .) (Potential).  
 FT CAROCHD 449 449 N-linked (GlcNAc. . .) (Potential).  
 FT CAROCHD 567 567 N-linked (GlcNAc. . .) (Potential).  
 FT CAROCHD 575 575 N-linked (GlcNAc. . .) (Potential).  
 FT CAROCHD 631 631 N-linked (GlcNAc. . .) (Potential).  
 FT CAROCHD 673 673 N-linked (GlcNAc. . .) (Potential).  
 FT CAROCHD 713 713 N-linked (GlcNAc. . .) (Potential).  
 FT CAROCHD 812 812 N-linked (GlcNAc. . .) (Potential).  
 FT CAROCHD 868 868 N-linked (GlcNAc. . .) (Potential).  
 FT CAROCHD 890 890 N-linked (GlcNAc. . .) (Potential).  
 FT CAROCHD 990 990 N-linked (GlcNAc. . .) (Potential).  
 SQ SEQUENCE 1621 AA; 174919 MW; 168528F21AABE22 CRC64;

Query Match 100.0%; Score 47; DB 1; Length 1621;  
 Best Local Similarity 100.0%; Pred. No. 2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDDLHV 10  
 DB 1224 SLAMDDLHV 1233

RESULT 5  
 ID PTB1\_SCHPO STANDARD; PRT; 311 AA.  
 AC P46360;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Type II protein geranyl-geranyltransferase beta subunit (EC 2.5.1.60)  
 DE (Type II protein geranyl-geranyltransferase beta subunit) (CGHase-II-beta)  
 GN Name=ptb1; ORFNames=SPAC167.02;  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OC NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96310627; PubMed=8740421;  
 RA Godfrey R., Davey J.;  
 RT "Sequence of ptb1, a gene for the beta subunit of the type-II  
 RT geranylgeranyltransferase from the fission yeast Schizosaccharomycetes  
 RT pombe.";  
 RL yeast 12:479-482(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21849401; PubMed=11859360; DOI=10.1038/nature724;  
 RA Wood V., Gwilliam R., Rajandream A.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Murphy K., Murphy L., Nisbet D., Odeli C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders R., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymptz B.,  
 RA Weljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Leinach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaie V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucifora M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,

RA Cerutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomycetes pombe.";  
 RL Nature 415:871-880(2002).  
 CC -1- FUNCTION: Catalyzes the transfer of a geranyl-geranyl moiety from  
 CC geranyl-geranyl pyrophosphate to proteins having the C-terminal  
 CC -XCC or -XCCX, where both cysteines may become modified.  
 CC -1- CATALYTIC ACTIVITY: 2 geranylgeranyl diphosphate + protein-  
 CC -1- CYSTEINE = 2 S-geranylgeranyl-protein + 2 diphosphate.  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit.  
 CC -1- SIMILARITY: Belongs to the protein prenyltransferase beta subunit  
 CC family.  
 CC -1- SIMILARITY: Contains 5 PFTB repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: X92183; CA63094.1; -  
 CC EMBL: A035248; CA22847.1; -  
 CC PIR: S65091; S65091.  
 CC HSBP; Q08603; LITX.  
 CC DR GenesB\_Spombe; SPAC167.02; -  
 CC DR InterPro; IPR001330; Prenyltrans.  
 CC DR InterPro; IPR008930; Terp\_cyc\_toroid.  
 CC DR Pfam; PF00432; Prenyltrans; 5.  
 CC KW Prenyltransferase; Repeat; Transferase; Zinc.  
 FT REPEAT 54 95 PFTB 1.  
 FT REPEAT 102 143 PFTB 2.  
 FT REPEAT 150 191 PFTB 3.  
 FT REPEAT 197 239 PFTB 4.  
 FT REPEAT 246 288 PFTB 5.  
 FT METAL 224 224 Zinc (By similarity).  
 FT METAL 226 226 Zinc (By similarity).  
 FT METAL 276 276 Zinc (By similarity).  
 SQ SEQUENCE 311 AA; 35092 MW; 2AB617FC769D3B08 CRC64;

Query Match 78.7%; Score 37; DB 1; Length 311;  
 Best Local Similarity 88.9%; Pred. No. 43;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMDDLHV 10  
 DB 91 LAMDDLHV 99

RESULT 6  
 ID Q8EC13 PRELIMINARY; PRT; 792 AA.  
 AC Q8EC13;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Hypothetical protein S03159.  
 GN OrderedLocustNames=S03159;  
 OS Shewanella oneidensis.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;  
 OC Alteromonadaceae; Shewanella.  
 CC NCBI\_TaxID=70863;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NR-1.  
 RX MEDLINE=22297686; PubMed=12368813; DOI=10.1038/nbt749;  
 RA Heidelberg J.F., Paulsen I.T., Nelson K.B., Gaidos E.J., Nelson W.C.,  
 RA Reed T.D., Eisen J.A., Seshadri R., Ward N.J., Mehta B.A.,  
 RA Clayton R.A., Meyer J.A., Teapin A., Scott J., Beaman M.J.,  
 RA Brinkac L.M., Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S.,  
 RA Hett D.H., Kolonay J.F., Madupu R., Peterson J.D., Umayam L.A.,

RA White O., Wolf A.M., Vamathavan J.J., Weidman J.F., Impirain M.,  
 RA Lee K., Berry K.O., Lee C., Mueller J., Knout H.M., Gill J.,  
 RA Ueberback T.R., McDonald L.A., Feldblum T.V., Smith H.O.,  
 RA Venter J.C., Neilson K.H., Fraser C.M.,  
 RT "Genome sequence of the dissimilatory metal ion-reducing bacterium  
 RT Shewanella oneidensis",  
 RL Nat. Biotechnol. 20:1118-1123(2002).  
 DR EMBL: AE015753; AAN56159.1; -.  
 DR TIGR: SC01559; -.  
 DR InterPro: IPR010344; DUF940.  
 DR Pfam: PF06082; DUF940; 1.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 792 AA; 88982 MW; F93CEFA5D24966B CRC64;

Query Match 78.7%; Score 37; DB 2; Length 792;  
 Best Local Similarity 80.0%; Pred. No. 1;e+02;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
 DB 18 SLAMDLHLV 27

RESULT 7  
 Q8UJFV0 PRELIMINARY; PRT; 1161 AA.  
 AC Q8UJFV0;  
 DT 01-OCT-2002 (TRENBLREL. 22, Created)  
 DT 01-OCT-2002 (TRENBLREL. 22, Last sequence update)  
 DT 01-MAR-2004 (TRENBLREL. 26, Last annotation update)  
 DE SI:42107016.1 (Novel protein similar to vertebrate anaplastic lymphoma  
 kinase (ALK) and leukocyte tyrosine kinase receptor precursor (LTK or  
 TYK) (Fragment).  
 CN Name=SI:42107016.1;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OC NCBI\_TaxID=7955;  
 RX SEQUENCE FROM N.A.  
 RA Hammett S.;  
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL596022; CQ43463.1; -.  
 DR HSP; G62838; IUDF.  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO: GO:0004872; F:receptor activity; IEA.  
 DR GO: GO:0016740; F:transferase activity; IEA.  
 DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR011009; Kinase\_like  
 DR InterPro: IPR002172; LDI\_receptor\_A.  
 DR InterPro: IPR000719; Prot\_Kinase.  
 DR InterPro: IPR002011; ReceptTyKinsit.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR InterPro: IPR008266; Tyr\_kinase\_AS.  
 DR Pfam: PF00057; Ldi\_recept\_a; 1.  
 DR Pfam: PF00069; Pkinase; 1.  
 DR PRINTS: PR00109; TYRKINASE  
 DR PRODOM: PD000001; Prot\_Kinase; 1.  
 DR SMART; SM00192; Ldlr; 1.  
 DR SMART; SM00219; Tyrc; 1.  
 DR PROSITE; PS01209; LdlrA\_1; UNKNOWN\_1.  
 DR PROSITE; PS00068; LdlrA\_2; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00039; RECEPTOR\_TYR\_KIN\_II; UNKNOWN\_1.  
 KW ATP-binding; kinase; Receptor; Transferase; Tyrosine-protein kinase.  
 FT NON\_TER 1  
 SQ SEQUENCE 1161 AA; 124851 MW; 601210E788B9AA41 CRC64;

Query Match 78.7%; Score 37; DB 2; Length 1161;

Best Local Similarity 70.0%; Pred. No. 1.7e+02;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
 DB 670 SLAMDLHLV 679

RESULT 8  
 Q9XHX1 PRELIMINARY; PRT; 376 AA.  
 ID Q9XHX1;  
 AC Q9XHX1;  
 DT 01-NOV-1999 (TRENBLREL. 12, Created)  
 DT 01-NOV-1999 (TRENBLREL. 12, Last sequence update)  
 DT 01-OCT-2004 (TRENBLREL. 28, Last annotation update)  
 DE Putative geranylgeranyl pyrophosphate synthase.  
 CN Name=OSJNB0049820.22; Synonyms=PG013C11.25;  
 OS Oryza sativa (Japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Erihartoideae; Oryzoideae; Oryza.  
 CN NCBI\_TaxID=39947;  
 RX SEQUENCE FROM N.A.  
 RA Buell R., Bentol M.-I., Lin X., Mason T.M., Unayam L., Shea T.P.,  
 RA Fujii C.Y., Shen M., Fraser C.M.,  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,  
 RA Wu J., Nishimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,  
 RA Hosokawa S., Masuoka M., Arikawa K., Chiden Y., Hayashi M.,  
 RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,  
 RA Hijikata S., Honda M., Ichikawa Y., Iiduma A., Iijima M., Ikeda M.,  
 RA Ikeno M., Itoh S., Itoh T., Itoh Y., Iwabuuchi A., Kamiya K.,  
 RA Katsura W., Katagiri T., Kikuta A., Kobayashi N., Kono I.,  
 RA Kuchita K., Maehara T., Mizuno H., Mizudayashi T., Nakai Y.,  
 RA Nagasaki N., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,  
 RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,  
 RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terashima K., Tsuji K.,  
 RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshitake R., Yukawa K.,  
 RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,  
 RA Yano M., Jiang J., Gojobori T.,  
 RT "The genome sequence and structure of rice chromosome 1.",  
 RL Nature 420:312-316(2002).  
 DR EMBL: AC007789; AAD38295.1; -.  
 DR EMBL: AP002865; BAB18334.1; -.  
 DR Gramene; Q9XHX1; -.  
 DR GO: GO:0008299; P:isoprenoid biosynthesis; IEA.  
 DR InterPro: IPR000092; Polyprenyl\_synth.  
 DR InterPro: IPR008949; Terpenoid\_synth.  
 DR Pfam: PF00348; polyprenyl\_synth\_1.  
 DR PROSITE; PS00723; POLYPRENYL\_SYNTHET\_1; 1.  
 DR PROSITE; PS00444; POLYPRENYL\_SYNTHET\_2; 1.  
 SQ SEQUENCE 376 AA; 39575 MW; F737BFB5B8B3B426 CRC64;

Query Match 76.6%; Score 36; DB 2; Length 376;  
 Best Local Similarity 70.0%; Pred. No. 84;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
 DB 250 SLAMDLHLV 259

RESULT 9  
 ID Q7SBR0 PRELIMINARY; PRT; 461 AA.  
 AC Q7SBR0;  
 DT 05-JUN-2004 (TRENBLREL. 27, Created)  
 DT 05-JUN-2004 (TRENBLREL. 27, Last sequence update)  
 DT 05-JUN-2004 (TRENBLREL. 27, Last annotation update)  
 DE ACR271CP.

```

GN Name=ACR271C; (Yeast) (Eremothecium gossypii).
OS Ashbya gossypii (Yeast); Ascomycota; Saccharomycotina; Saccharomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OC NCB1_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10895;
RX PubMed=15001715;
RA Dietrich F.S., Voegel S., Brachat S., Lerch A., Gares K., Steiner S.,
  Mohr C., Pohlmann R., Luedi P., Choi S., Wing R.A., Flavier A.,
  Rafferty T.D., Philippson P.,
  "The Ashbya gossypii genome as a tool for mapping the ancient
  Saccharomyces cerevisiae genome."
  Science 304:304-307(2004).
  EMBL: AE016888; AAS51497.1;
  DR AGD; ACR271C;
  SQ SEQUENCE 461 AA; 52404 MW; 5D6FDBF10928C846 CRC64;

Query Match
Best Local Similarity 76.6%; Score 36; DB 2; Length 461;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 80 SLAMDLLH 9
1 SLAMDLLH 9
PRT; 461 AA.
AAS51497; PRELIMINARY;
AC AAS51497; (TREMBLrel. 27, Created)
DT 23-APR-2004 (TREMBLrel. 27, Last sequence update)
RT 23-APR-2004 (TREMBLrel. 27, Last annotation update)
DE ACR271C.
GN ACR271C.
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetidae; Saccharomycetaceae; Eremothecium.
ON NCB1_TaxID=33169;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10895;
RX PubMed=15001715;
RA Dietrich F.S., Voegel S., Brachat S., Lerch A., Gares K., Steiner S.,
  Mohr C., Pohlmann R., Luedi P., Choi S., Wing R.A., Flavier A.,
  Gaffney T.D., Philippson P.,
  "The Ashbya gossypii genome as a tool for mapping the ancient
  Saccharomyces cerevisiae genome."
  Science 304:304-307(2004).
  EMBL: AE016888; AAS51497.1;
  DR EMBL; AE016888; AAS51497.1;
  SQ SEQUENCE 461 AA; 52404 MW; 5D6FDBF10928C846 CRC64;

Query Match
Best Local Similarity 76.6%; Score 36; DB 2; Length 461;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 80 SLAMDLLH 9
1 SLAMDLLH 9
PRT; 763 AA.
AAS51497; PRELIMINARY;
AC AAS51497; (TREMBLrel. 26, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
RT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=NCU08177.1;
OS Neurospora crassa.

```

```

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OC NCB1_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
  Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Reiman B.,
  Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
  Ou D., Iankiev P., Pedersen D., Nelson M., Mashburne M.,
  Seltrennikoff C.P., Kisey J.A., Braun E.L., Zelter A., Schulte U.,
  Koche G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
  Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
  Kamil M., Kamysheles M., Maccell E., Blake C., Rudd S., Frisman D.,
  Krysstova S., Rasmussen C., Metzger R.L., Perkins D.V., Kroken S.,
  Cogoni C., Macho G., Catchside D., Li W., Pratt R.V., Osmani S.A.,
  Desouza C.C., Glaes L., Orbach M.J., Berglund J., Voelker R.,
  Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
  Natvig D.O., Alex L.A., Manhaupt G., Ebbole D.J., Freitag M.,
  Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.,
  "The Genome Sequence of the Filamentous Fungus Neurospora crassa."
  Nature 010-012003.
  RL Nature 010-012003.
  CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
  preliminary data.
  CC EMBL; AABX0100370; EAA30334.1;
  DR GO: GO:0005524; F:ATP binding; IEA.
  DR GO: GO:0004672; F:Protein kinase activity; IEA.
  DR GO: GO:0016740; F:transferase activity; IEA.
  DR GO: GO:0006468; F:Protein amino acid phosphorylation; IEA.
  DR InterPro: IPR007719; F:Protein kinase.
  DR InterPro: IPR008271; Ser_Thr_Pkin_AS.
  DR Pfam: PF00653; Pkinase; 1.
  DR PRODOM: PD000001; Prot_kinase; 1.
  DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
  DR PROSITE: PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
  KW Hypothetical protein.
  SQ SEQUENCE 763 AA; 84728 MW; 414B0DE9089D24D5 CRC64;

Query Match
Best Local Similarity 77.8%; Score 36; DB 2; Length 763;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 642 ALAIDLH 650
1 ALAIDLH 9
PRT; 872 AA.
Q8PUW3; PRELIMINARY;
AC Q8PUW3; (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
RT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE ATP-dependent DNA ligase.
GN Name=1193; Order=locusNames=XAC2414;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
ON NCB1_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XY 101;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
  Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
  Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo J.B.A.,
  Amaral G., Camarvan F., Cardoso J., Chamberg F., Ciapina L.P.,
  Ciccarelli R.M.B., Coutinho L.L., Curisio-Santos J.R., El-Dorri H.,
  Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Garber A.,
  Formighieri E.F., Franco M.C., Greggio C.C., Gunter A.,
  Katsuyama A.M., Kishi L.T., Leite K.P., Lemos E.G.M., Lemos M.V.F.,

```

RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
 RA Moreira L.M., Novo M.T.W., Okura V.K., Oliveira M.C., Oliveira V.R.,  
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,  
 RA Spindola L.A.F., Takita M.A., Tamura R.B., Teixeira E.C., Tezza R.I.D.,  
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,  
 RA Seubal J.C., Kitajima J.P.,  
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
 RT host specificities.",  
 RI Nature 417:459-463(2002).  
 CC -1- CATALYTIC ACTIVITY: ATP + (deoxyribonucleotide) (N) +  
 CC {deoxyribonucleotide} (M) = AMP + diphosphate +  
 CC {deoxyribonucleotide} (N+M).  
 CC -1- SIMILARITY: Belongs to the ATP-dependent DNA ligase family.  
 DR EMBL: AE01878; AAC37266.1; -.  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0003910; F:DNA ligase (ATP) activity; IEA.  
 DR GO: GO:0016874; F:ligase activity; IEA.  
 DR GO: GO:0003676; F:nucleic acid binding; IEA.  
 DR GO: GO:0006310; P:DNA recombination; IEA.  
 DR GO: GO:0006281; P:DNA repair; IEA.  
 DR GO: GO:0006260; P:DNA replication; IEA.  
 DR InterPro: IPR009377; DNA\_ligase.  
 DR InterPro: IPR008994; Nucleic\_acid\_OB.  
 DR Pfam: PF04679; DNA\_ligase\_A\_C; 1.  
 DR Pfam: PF01068; DNA\_ligase\_A\_M; 1.  
 DR PROSITE: PS00697; DNA\_LIGASE\_A1; UNKNOWN\_1.  
 DR PROSITE: PS01650; DNA\_LIGASE\_A3; 1.  
 KM ATP-binding; Complete proteome; DNA recombination; DNA repair;  
 KM DNA replication; Ligase.  
 SQ SEQUENCE 872 AA; 96305 MW; 3C5AC9FBA2F08921 CRC64;

Query Match 76.6%; Score 36; DB 2; Length 872;  
 Best Local Similarity 70.0%; Pred. No. 2e+02;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLIH 10  
 Db 325 ALAVFDLHV 334

RESULT 13

OC 023723 PRELIMINARY; PRT; 989 AA.  
 AC 023723;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative arabinose kinase.  
 GN Name=ISA1;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RX MEDLINE=98192529; PubMed=9524266;  
 RA Cy 1., Aubourg S., Sheron S., Cobbett C.S., Cheron A., Kreis M.,  
 RA Lecharny A.,  
 RT "Analysis of a 14-kb fragment containing a putative cell wall gene and  
 RT a candidate for the ARA1, arabinose kinase, gene from chromosome IV of  
 RT Arabidopsis thaliana.",  
 RI Gene 209:201-210(1998).  
 DR EMBL: Y14404; CAA4753.1; -.  
 DR GO: GO:0005737; C:cyclopiaem; IEA.  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0016301; F:kinase activity; IEA.  
 DR GO: GO:0016773; F:phosphotransferase activity; IEA.  
 DR GO: GO:0016740; F:transferase activity; IEA.  
 DR GO: GO:0008152; P:metabolism; IEA.  
 DR GO: GO:0016310; P:phosphorylation; IEA.  
 DR InterPro: IPR006204; GHMP\_kinase.

DR InterPro: IPR006206; Mey\_galkinase.  
 DR Pfam: PF00288; GHMP\_kinases; 1.  
 DR PRINTS: PR00959; MEVGALKINASE.  
 KM ATP-binding; Kinase; Transferase.  
 SQ SEQUENCE 989 AA; 10846 MW; CC954177C888438B CRC64;

Query Match 76.6%; Score 36; DB 2; Length 989;  
 Best Local Similarity 77.8%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLIH 9  
 Db 464 SLAMDLIH 472

RESULT 14

OC 023461 PRELIMINARY; PRT; 1039 AA.  
 AC 023461;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Galactokinase like protein.  
 GN Name=d14105w; Synonyms=ATG16130;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RX SEQUENCE FROM N.A.  
 RA Bevan M., Stiekema W., Murphy G., Wambutt R., Pohl T., Terryn N.,  
 RA Kreis M., Kavanagh T., Entian K.D., Rieger W., James R.,  
 RA Pulgadennech P., Hatzopoulos P., Obermaier B., Duesterhoft A.,  
 RA Jones J., Palme K., Ansgore W., Delseny M., Bancroft I., Mewes H.W.,  
 RA Schueller C., Chalmers N.,  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RX SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RX SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: Z97340; CAB10392.1; -.  
 DR EMBL: AL161543; CAB8655.1; -.  
 DR PIR: F71427; F71427.  
 DR GO: GO:0005737; C:cyclopiaem; IEA.  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0016301; F:kinase activity; IEA.  
 DR GO: GO:0016773; F:phosphotransferase activity; IEA.  
 DR GO: GO:0016740; F:transferase activity; IEA.  
 DR GO: GO:0008152; P:metabolism; IEA.  
 DR GO: GO:0016310; P:phosphorylation; IEA.  
 DR InterPro: IPR006204; GHMP\_kinase.  
 DR InterPro: IPR006206; Mey\_galkinase.  
 DR Pfam: PF00288; GHMP\_kinases; 1.  
 DR PRINTS: PR00959; MEVGALKINASE.  
 KM ATP-binding; Kinase; Transferase.  
 SQ SEQUENCE 1039 AA; 114260 MW; 82843127FC866B5 CRC64;

Query Match 76.6%; Score 36; DB 2; Length 1039;  
 Best Local Similarity 77.8%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLIH 9  
 Db 514 SLAMDLIH 522

RESULT 15

Q7GRY1

```

ID Q7ORY1 PRELIMINARY; PRT; 1066 AA.
AC Q7ORY1;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE GIP_549_14267.17467
OS Giardia lamblia ATCC 50803.
OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.
OX NCBI_TaxID=184922;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W3 C6;
RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
RT Olsen G.O., Sogin M.L.; Giardia lamblia genome.
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
CC EMBL, AACB0100127; EAA37782.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:Protein kinase activity; IEA.
DR GO; GO:0016740; F:Transferase activity; IEA.
DR GO; GO:0006468; P:Protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00069; Kinase; 1.
DR PRODOM; PD000001; Prot_kinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
SQ SEQUENCE 1066 AA; 11953 MW; E214831E6780AE27 CRC64;

Query Match 76.6%; Score 36; DB 2; Length 1066;
Best Local Similarity 77.8%; Pred. No. 2.5e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

CY 2 LAMDDLHV 10
|:|||||
Db 236 LGLDDLHV 244

```

Search completed: November 30, 2004, 07:17:26  
 Job time : 78.4103 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 ; Search time 77.1795 Seconds

(without alignments)  
46,480 Million cell updates/sec

Title: US-10-008-377A-1

Perfect score: 47

Sequence: 1 SLAMDLMHV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 35672929 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A\_Geneseq\_23sep04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	100.0	10	AA022983	AA022983 Human p28
2	47	100.0	45	AB073862	AB073862 Human ALK
3	47	100.0	45	AB074476	AB074476 Tyrosine
4	47	100.0	525	AA078521	AA078521 NPM/ALK f
5	47	100.0	525	AB004368	AB004368 Human exp
6	47	100.0	572	AA078520	AA078520 Human exp
7	47	100.0	675	AB020950	AB020950 Partial A
8	47	100.0	680	AA073858	AA073858 Novel hum
9	47	100.0	680	AB073858	AB073858 Human NFM
10	47	100.0	680	AB095067	AB095067 Human tpa
11	47	100.0	680	AB004347	AB004347 Human exp
12	47	100.0	680	AB004362	AB004362 Human exp
13	47	100.0	680	AB004358	AB004358 Human exp
14	47	100.0	680	AB074472	AB074472 NPM/ALK f
15	47	100.0	1620	AA018498	AA018498 Human ALK
16	47	100.0	1620	AB074471	AB074471 Human ins
17	43	91.5	9	AA022985	AA022985 Human p28
18	39	83.0	9	AA022986	AA022986 Human p28
19	36	76.6	376	ADM98606	ADM98606 Human p28
20	36	76.6	376	ADM98621	ADM98621 Geranylge
21	36	76.6	376	ADM98647	ADM98647 Geranylge
22	35	74.5	1039	AB092855	AB092855 Herdida
23	35	74.5	187	AB070368	AB070368 Novel hum
24	35	74.5	302	AB025901	AB025901 Aspergill
25	35	74.5	334	AB026501	AB026501 Aspergill

26	35	74.5	426	8	ADH22360	Adh22360 Human rec
27	35	74.5	436	8	ADH34619	Adh34619 Endothel
28	35	74.5	441	2	AAR25394	Aar25394 Sequence
29	35	74.5	442	2	AAR30886	Aar30886 Etb recep
30	35	74.5	442	4	ABB56352	Abb56352 Non-endog
31	35	74.5	442	5	ABR57085	Abbr57085 Mouse isc
32	35	74.5	442	6	ABR58526	Abbr58526 Human end
33	35	74.5	442	6	ABG74670	Abg74670 Human EDN
34	35	74.5	442	6	ABP81815	Abp81815 Human end
35	35	74.5	442	6	ABU61640	Abu61640 Human end
36	35	74.5	442	7	ADE61732	Ad61732 Human pro
37	35	74.5	442	7	ADE61730	Ad61730 Rat Prote
38	35	74.5	442	7	ADE61736	Ad61736 Human Pro
39	35	74.5	442	7	ADD46389	Ad46389 Rat Prote
40	35	74.5	442	7	ADD46391	Ad46391 Human Pro
41	35	74.5	442	7	AD339949	Ad339949 Human HUE
42	35	74.5	442	7	ADO29303	Ado29303 Human GPC
43	35	74.5	442	8	ADO29304	Ado29304 Mouse GPC
44	35	74.5	442	8	ADO18283	Ado18283 Human sof

## ALIGNMENTS

RESULT 1	AA022983	standard; peptide; 10 AA.
ID	AA022983	
XX	AA022983	
AC	AA022983	
XX	17-SEP-2003	(first entry)
DT	17-SEP-2003	
DE	Human p280-89 ALK-derived HLA-A*0201 restricted CTL epitope peptide.	
HLA-A*0201 binding; anaplastic lymphoma kinase; ALK; CTL epitope; ALCU;		
KW	cytotoxic T-lymphocyte; receptor tyrosine; NPM/ALK fusion protein; human;	
KW	oncogenic; anaplastic large cell lymphoma; neuroblastoma; gene therapy;	
KW	t(2;5) translocation; neoplasia; cytostatic; immunotherapy; alkoma;	
KW	p280-89.	
OS	Homo sapiens.	
XX		
PN	W02003042243-A2.	
XX		
PD	22-MAY-2003.	
XX		
PF	14-NOV-2002; 2002MO-EP012764.	
XX		
PR	15-NOV-2001; 2001US-00008377.	
XX		
PA	(MAST-) 1ST NAZ STUDIO E CURA DEI TUMORI.	
PA	(SUD-) 1ST JUDE CHILDREN'S RES HOSPITAL.	
XX		
PI	Gambacorti-Passerini C, Passoni L;	
DR	WPI; 2003-441791/41.	
XX		
PT	New HLA-A*0201-binding Anaplastic Lymphoma Kinase (ALK) peptide,	
PT	useful for preparing a composition for treating ALK-positive lymphoma,	
PT	neuroblastoma or ALK-expressing neoplasia.	
XX		
PS	Claim 1; Page 3; 33pp; English.	
XX		
CC	The invention relates to a novel HLA-A*0201-binding anaplastic lymphoma	
CC	kinase (ALK)-derived peptide, of 9-10 amino acids, which may act as a	
CC	cytotoxic T-lymphocyte (CTL)-epitope. ALK is a receptor tyrosine kinase	
CC	which forms part of the NPM/ALK fusion protein (oncogenic lymphoma kinase	
CC	fusion protein) in ALCU (anaplastic large cell lymphoma)-derived cell	
CC	lines. More than 50% of ALCU cases possess a t(2;5) chromosomal	
CC	translocation that leads to the expression of the NPM/ALK fusion protein	
CC	which forms a potent oncogene when constitutively activated. Translocated	
CC	ALK is a widely expressed tumour-associated antigen characteristic of ALK	

CC -positive lymphomas, neuroblastomas and ALK-expressing neoplasias. The  
 CC peptides of the invention demonstrate cytosstatic activity and induce an  
 CC MHC (major histocompatibility complex) class I restricted cytotoxic  
 CC lymphocyte response against tumour cells expressing the NPM/ALK fusion  
 CC protein. Hence, the peptides may be utilised during the treatment, via  
 CC immunotherapy, of ALK-positive lymphomas such as alkomas, neuroblastomas  
 CC and ALK-expressing neoplasias. In addition, the peptides may be used  
 CC during gene therapy. The current sequence is that of the human p280-89  
 CC ALK-derived HLA-A\*0201 restricted CTL epitope peptide of the invention  
 XX  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 47; DB 6; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLAMDILHY 10  
 DB 1 SLAMDILHY 10

## RESULT 2

AA873862 standard; peptide; 45 AA.

AA873862;

15-MAY-2001 (first entry)

Human ALK catalytic domain sequence #3.

Human; anaplastic lymphoma kinase; ALK; nucleophosmin; NPM;

chromosomal translocation; cancer; NPM/ALK; catalytic domain.

Homo sapiens.

US6174674-B1.

16-JAN-2001.

19-JUN-1998; 98US-00100089.

03-DEC-1993; 93US-00160861.

12-OCT-1995; 95US-00542363.

(SUUD-) ST JUDE CHILDREN'S RES HOSPITAL.

Morris SW, Look AT;

WPI; 2001-243208/25.

DR Detection of chromosomal rearrangement or translocations present in t(2;  
 PT 5) for cancer diagnosis comprises using and identifying anaplastic  
 PT lymphoma kinase and nucleophosmin genes in nucleic acid hybridization  
 techniques.

Example 2; Fig 2C; 87pb; English.

CC The present sequence is given in a specification relating to a method for  
 CC detecting a chromosomal rearrangement involving a breakpoint in the  
 CC anaplastic lymphoma kinase (ALK) or nucleophosmin (NPM) gene. The method  
 CC is based on the identification and sequence determination of a novel  
 CC gene, ALK, which is fused to the gene encoding NPM in translocations  
 CC present in t(2;5) lymphoma cells. The method is useful for detecting a  
 CC t(2;5) chromosomal rearrangement, or a chromosomal rearrangement  
 CC involving the ALK gene or NPM gene. It is particularly useful in  
 CC determining if particular cells or tissues express ALK or NPM/ALK coding  
 CC sequences. The method is also useful in diagnostic assays to determine,  
 CC for example, if a mammal has cancer or a genetic predisposition to cancer  
 XX  
 SQ Sequence 45 AA;

Query Match 100.0%; Score 47; DB 4; Length 45;

Best Local Similarity 100.0%; Pred. No. 0.069;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLAMDILHY 10  
 DB 5 SLAMDILHY 14

## RESULT 3

ABG74476 standard; peptide; 45 AA.

ABG74476;

11-APR-2003 (first entry)

Tyrosine kinase insulin receptor ALK fragment SEQ ID 14.

ALK; human; nucleophosmin; NPM; anaplastic lymphoma kinase; chromosome 5;

t(2; 5) lymphoma cell; chromosomal rearrangement; chromosome 2;

nucleolar phosphoprotein; centromeric; telomeric.

Homo sapiens.

US6451997-B1.

28-SEP-2000; 2000US-00670827.

03-DEC-1993; 93US-00160861.

12-OCT-1995; 95US-00542363.

19-JUN-1998; 98US-00100089.

(SUUD-) ST JUDE CHILDREN'S RES HOSPITAL.

Morris SW, Look AT;

WPI; 2003-101739/09.

PT Kit for use in method of detecting t(2;5) chromosomal rearrangements or  
 PT rearrangements involving nucleolar phosphoprotein (NPM) or anaplastic  
 PT lymphoma kinase (ALK) comprises probes that hybridizes with NPM and ALK  
 PT genes.

Example 1; Fig 2C; 87pb; English.

CC This invention describes a novel kit for use in a method of detecting t(2;  
 CC ;5) chromosomal rearrangements (CR) or CRs involving nucleolar  
 CC phosphoprotein (NPM) and anaplastic lymphoma kinase (ALK). The method  
 CC comprises labelled probes that hybridize to a sequence of 2442 or 6226  
 CC nucleotides or their complement, and to region of human chromosome 5/2  
 CC having wild-type NPM/ALK genes (the region is centromeric/telomeric to  
 CC nucleolus). This sequence represents a polypeptide sequence described in  
 CC the disclosure of the invention

Sequence 45 AA;

Query Match 100.0%; Score 47; DB 6; Length 45;  
 Best Local Similarity 100.0%; Pred. No. 0.069;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLAMDILHY 10  
 DB 5 SLAMDILHY 14

## RESULT 4

AA878521 standard; protein; 525 AA.

AA878521;

DT 25-MAR-2003 (revised)  
 XX 15-NOV-1995 (first entry)  
 DE NPM/ALK fusion protein.  
 XX  
 XX  
 KW ALK protein; protein-tyrosine-kinase; fusion protein; NPM protein;  
 KW nuclear phosphoprotein; anaplastic large cell lymphoma; t(2;  
 XX 5) translocation; diagnosis; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO9515331-A1.  
 XX  
 PD 08-JUN-1995.  
 XX  
 XX 05-DEC-1994; 94WO-US013947.  
 XX  
 XX 03-DEC-1993; 93US-00160861.  
 XX  
 PA (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 XX  
 PI Morris SM, Look AT;  
 XX  
 XX WPI, 1995-215226/28.  
 XX  
 DR N-P8DB; AAQ95542.  
 XX  
 PT Methods for detecting human t(2;5) lymphoma - for detection and diagnosis  
 PT of anaplastic large cell lymphoma(s).  
 XX  
 PS Claim 16; Page 42-44; 70pp; English.  
 XX  
 CC The translocation event that occurs in human t(2;5) lymphoma brings  
 CC sequences from the nucleolar phosphoprotein (NPM) gene on chromosome 5q35  
 CC to those from a protein-tyrosine-kinase (ALK) gene on chromosome-2q23.  
 CC Nucleic acids encoding a partial ALK protein (AAR78520) and the ALK/NPM  
 CC fusion protein (AAR78521) were isolated (AAQ95541, AAQ95542).  
 CC Identification of the NPM/ALK fusion allows lymphoma diagnosis, or  
 CC therapy through the use of antisense RNA or ribozymes. (Updated on 25-MAR  
 CC 2003 to correct PN field.)  
 XX  
 SQ Sequence 525 AA;  
 XX  
 Query Match 100.0%; Score 47; DB 2; Length 525;  
 Best Local Similarity 100.0%; Pred. No. 0.99;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 QY 1 SLAMF1LHV 10  
 DB 280 SLAMF1LHV 289  
 XX  
 RESULT 5  
 ABU04368  
 ID ABU04368 standard; protein; 525 AA.  
 XX  
 XX ABU04368;  
 AC  
 XX 29-JAN-2003 (first entry)  
 DT  
 XX  
 XX Human expressed protein tag (EPT) #1034.  
 DE  
 XX  
 XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
 KW protease; protein inhibitor; transporter; cytoskeletal protein;  
 KW receptor; transcription factor; cancer; MHC; colon cancer; gastric cancer;  
 KW major histocompatibility complex; myeloma; leukemia; leukemia.  
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX WO200278524-A2.  
 XX  
 XX 10-OCT-2002.  
 PD  
 XX

EF 28-MAR-2002; 2002WO-US009671.  
 XX  
 XX 28-MAR-2001; 2001US-0279495P.  
 PR 21-MAY-2001; 2001US-0292544P.  
 PR 08-AUG-2001; 2001US-0310801P.  
 PR 01-OCT-2001; 2001US-0326370P.  
 PR 04-DEC-2001; 2001US-0336780P.  
 PR 20-FEB-2002; 2002US-0358985P.  
 XX  
 XX (ZYCO-) ZYCO INC.  
 XX  
 XX Chicx RM, Tomlinson AJ, Urban RG;  
 XX  
 XX WPI, 2003-040607/03.  
 XX  
 XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 PT cytoskeletal proteins, receptors or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PT leukemia.  
 XX  
 XX Example 2; SEQ ID NO 1034; 134pp; English.  
 XX  
 CC The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide. The purified polypeptide, or the antibody that binds to this  
 CC polypeptide, is useful for treating cancer. The polypeptide is also  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptides and  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
 CC lymphoma or leukemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: This sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 525 AA;  
 XX  
 Query Match 100.0%; Score 47; DB 6; Length 525;  
 Best Local Similarity 100.0%; Pred. No. 0.99;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 QY 1 SLAMF1LHV 10  
 DB 280 SLAMF1LHV 289  
 XX  
 RESULT 6  
 AAR78520  
 ID AAR78520 standard; protein; 572 AA.  
 XX  
 XX AAR78520;  
 AC  
 XX 25-MAR-2003 (revised)  
 DT 15-NOV-1995 (first entry)  
 XX  
 XX Partial ALK protein.  
 DE  
 XX  
 XX ALK protein; protein-tyrosine-kinase; fusion protein; NPM protein;  
 KW nuclear phosphoprotein; anaplastic large cell lymphoma; t(2;  
 KW 5) translocation; diagnosis; gene therapy.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX WO9515331-A1.  
 XX  
 XX 08-JUN-1995.  
 PD  
 XX 05-DEC-1994; 94WO-US013947.  
 PF

```

XX 03-DEC-1993; 93US-00160861.
XX (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
XX Morris SW, Look AT;
XX
XX WPI: 1995-215226/28.
XX N-PSDB; AAG95541.
XX
XX Methods for detecting human t(2;5) lymphoma - for detection and diagnosis
XX of anaplastic large cell lymphoma(s).
XX
XX Claim 21; Page 40-42; 70pp; English.
XX
XX The translocation event that occurs in human t(2;5) lymphoma brings
XX sequences from the nucleolar phosphoprotein (NPM) gene on chromosome 5q35
XX to those from a protein-tyrosine-kinase (ALK) gene on chromosome-2q23.
XX Nucleic acids encoding a partial ALK protein (AAR78520) and the ALK/NPM
XX fusion protein (AAR78521) were isolated (AAG95541, AAG95542).
XX Identification of the NPM/ALK fusion allows lymphoma diagnosis, or
XX therapy through the use of antisense RNA or ribozymes. (Updated on 25-MAR
XX -2003 to correct PN field.)
XX
XX Sequence 572 AA;
XX
XX Query Match 100.0%; Score 47; DB 2; Length 572;
XX Best Local Similarity 100.0%; Pred. No. 1.1;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 SLAMDLHLV 10
XX 327 SLAMDLHLV 336
XX
XX RESULT 7
XX AAG20950
XX ID AAG20950 standard; protein; 675 AA.
XX
XX AC AAG20950;
XX
XX DT 13-FEB-2002 (first entry)
XX
XX DE Novel human diagnostic protein #20941.
XX
XX KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX OS Homo sapiens.
XX
XX PN WO200175067-A2.
XX
XX PD 11-OCT-2001.
XX
XX PF 30-MAR-2001; 2001WO-US008631.
XX
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX
XX PA (HYSE-) HYSEQ INC.
XX
XX PI Dymnac RT, Liu C, Tang YT;
XX
XX DR WPI: 2001-639362/73.
XX DR N-PSDB; AAS85137.
XX
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX PS Claim 20; SEQ ID NO 51309; 103pp; English.

```

```

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptides and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAG0010-A830377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 675 AA;
XX
XX Query Match 100.0%; Score 47; DB 4; Length 675;
XX Best Local Similarity 100.0%; Pred. No. 1.3;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 SLAMDLHLV 10
XX 275 SLAMDLHLV 284
XX
XX RESULT 8
XX AAB73858
XX ID AAB73858 standard; protein; 680 AA.
XX
XX AC AAB73858;
XX
XX DT 15-MAY-2001 (first entry)
XX
XX DE Human NPM/ALK fusion protein.
XX
XX KM Human; anaplastic lymphoma kinase; ALK; nucleophosmin; NPM;
XX chromosomal translocation; cancer; NPM/ALK; fusion.
XX
XX OS Homo sapiens.
XX
XX PN US6174674-B1.
XX
XX PD 16-JAN-2001.
XX
XX PF 19-JUN-1998; 98US-00100089.
XX
XX PR 03-DEC-1993; 93US-00160861.
XX PR 12-OCT-1995; 95US-00542363.
XX
XX PA (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
XX PI Morris SW, Look AT;
XX
XX DR WPI: 2001-243208/25.
XX DR N-PSDB; AAF76868.
XX
XX PT Detection of chromosomal rearrangement or translocations present in t(2;
XX 5) for cancer diagnosis comprises using and identifying anaplastic
XX lymphoma kinase and nucleophosmin genes in nucleic acid hybridization
XX techniques.
XX
XX PS Claim 1; Fig 2A; 87pp; English.
XX
XX The present sequence is given in a specification relating to a method for
XX detecting a chromosomal rearrangement involving a breakpoint in the

```

CC anaplastic lymphoma kinase (ALK) or nucleophosmin (NPM) gene. The method  
 CC is based on the identification and sequence determination of a novel  
 CC gene, ALK, which is fused to the gene encoding NPM in translocations  
 CC present in t(2;5) lymphoma cells. The method is useful for detecting a  
 CC t(2;5) chromosomal rearrangement, or a chromosomal rearrangement  
 CC involving the ALK gene or NPM gene. It is particularly useful in  
 CC determining if particular cells or tissues express ALK or NPM/ALK coding  
 CC sequences. The method is also useful in diagnostic assays to determine,  
 CC for example, if a mammal has cancer or a genetic predisposition to cancer  
 XX

CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or  
 CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.  
 CC P53), or selectively treating cells expressing mutant protein or cellular  
 CC protein isoform in a patient heterozygous for (II). The method is useful  
 CC for treating a disease e.g. haematopoietic disorder such as T or B cell  
 CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, AML, NHL and CML,  
 CC or a disease characterised by a solid tumour such as papillary thyroid  
 CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and  
 CC synovial sarcoma. The method is also useful for treating viral  
 CC infections. This represents a protein encoded by the DNA sequence of a  
 CC chromosome aberration  
 XX

Query Match 100.0%; Score 47; DB 4; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SLAMDDLHV 10  
 Db 280 SLAMDDLHV 289

Query Match 100.0%; Score 47; DB 5; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SLAMDDLHV 10  
 Db 280 SLAMDDLHV 289

RESULT 9  
 ABG95067  
 ID ABG95067 standard; protein; 680 AA.  
 XX  
 AC AEG95067;  
 XX  
 DT 04-DEC-2002 (first entry)  
 XX  
 DE Human translocation (2; 5) (p23; q35) protein.  
 XX  
 KM Chromosome aberration; oncogenic fusion protein; cancer;  
 KM proliferative disease; cellular protein isoform; heat shock protein 90;  
 KM HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;  
 KM T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;  
 KM acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;  
 KM acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;  
 KM papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;  
 KM rhabdomyosarcoma; synovial sarcoma; viral infection.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200269900-A2.  
 XX  
 PD 12-SEP-2002.  
 XX  
 PF 01-MAR-2002; 2002MO-US006518.  
 XX  
 PR 01-MAR-2001; 2001US-027751P.  
 XX

RESULT 10  
 ABU04347  
 ID ABU04347 standard; protein; 680 AA.  
 XX  
 AC ABU04347;  
 XX  
 DT 29-JAN-2003 (first entry)  
 XX  
 DE Human expressed protein tag (EPT) #1013.  
 XX  
 KM Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
 KM protease; protease inhibitor; transporter; cytoskeletal protein;  
 KM receptor; transcription factor; cancer; WBC;  
 KM major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 KM adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200278524-A2.  
 XX  
 PD 10-OCT-2002.  
 XX  
 PF 28-MAR-2002; 2002MO-US009671.  
 XX  
 PR 28-MAR-2001; 2001US-0279495P.  
 PR 21-MAY-2001; 2001US-0292544P.  
 PR 08-AUG-2001; 2001US-0310801P.  
 PR 01-OCT-2001; 2001US-0326370P.  
 PR 04-DEC-2001; 2001US-0336780P.  
 PR 20-FEB-2002; 2002US-0358985P.  
 XX  
 PA (ZYCO-) ZYCOS INC.  
 XX  
 PI Chicx RM, Tomlinson AJ, Urban RG;  
 XX  
 DR WPI; 2003-040607/03.  
 XX

PT Treating genetically-defined disease associated with chromosomal  
 PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative  
 PT diseases, involves administering an inhibitor of heat shock protein 90.  
 XX  
 DR WPI; 2002-698710/75.  
 DR N-PSDB; ABS73246.  
 XX  
 PI Fritz LC, Burrows RJ;  
 XX  
 PA (CONF-) CONFORMA THERAPEUTICS CORP.  
 XX  
 DR WPI; 2002-698710/75.  
 DR N-PSDB; ABS73246.  
 XX  
 PD 12-SEP-2002.  
 XX  
 PF 01-MAR-2002; 2002MO-US006518.  
 XX  
 PR 01-MAR-2001; 2001US-027751P.  
 XX

PT Treating genetically-defined disease associated with chromosomal  
 PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative  
 PT diseases, involves administering an inhibitor of heat shock protein 90.  
 XX  
 DR WPI; 2003-040607/03.  
 XX  
 PI Chicx RM, Tomlinson AJ, Urban RG;  
 XX  
 DR WPI; 2003-040607/03.  
 XX  
 PA New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 PA cytoskeletal proteins, receptors or transcription factors), useful for  
 PA treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PA leukemia.  
 XX  
 PI Example 2; SEQ ID NO 1013; 134pp; English.  
 XX

CC The invention describes a method of treating genetically-defined disease  
 CC associated with chromosomal aberrations yielding oncogenic fusion  
 CC proteins (I), treating cancerous cells containing (I) in a heterogeneous  
 CC cell population, treating proliferative diseases associated with mutant  
 CC protein or cellular protein isoforms (II) dependent on heat shock protein  
 CC (HSP)-90, or selectively treating cells expressing (II) involving  
 CC administering HSP90-inhibitor. The method is useful for treating  
 CC genetically-defined disease with chromosomal aberration yielding  
 CC oncogenic fusion protein, treating cancerous cells containing fusion  
 CC protein in heterogeneous cell population, treating proliferative disease

CC The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide. The purified polypeptide, or the antibody that binds to this

CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 680 AA;

Query Match 100.0%; Score 47; DB 6; Length 680;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDDLHV 10  
DB 280 SLAMDDLHV 289

## RESULT 11

ABU04362  
ID ABU04362 standard; protein; 680 AA.

AC ABU04362;

DT 29-JAN-2003 (first entry)

DE Human expressed protein tag (EPT) #1028.

XX  
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
XX protease; protease inhibitor; transporter; cytoskeletal protein;  
XX receptor; transcription factor; cancer; MHC;  
XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX Homo sapiens.

XX WO200278524-A2.

XX 10-OCT-2002.

XX 28-MAR-2002; 2002WO-US009671.

XX 28-MAR-2001; 2001US-0279495P.

XX 21-MAY-2001; 2001US-0292544P.

XX 08-AUG-2001; 2001US-0310801P.

XX 01-OCT-2001; 2001US-0326370P.

XX 04-DEC-2001; 2001US-0336780P.

XX 20-FEB-2002; 2002US-0358985P.

XX (ZYCO-) ZYCOS INC.

XX Chicx RM, Tomlinson AJ, Urban RG;

XX WPI; 2003-040607/03.

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
XX cytoskeletal proteins, receptors or transcription factors), useful for  
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
XX leukemia.

XX Example 2; SEQ ID NO 1028; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a  
XX fragment of a kinase, phosphatase, protease, protease inhibitor,  
XX transporter, cytoskeletal protein, receptor or transcription factor. The  
XX polypeptide is useful as an immunogenic composition for eliciting in a  
XX mammal an immunogenic response directed against any of the purified  
XX polypeptide. The purified polypeptide, or the antibody that binds to this

CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 680 AA;

Query Match 100.0%; Score 47; DB 6; Length 680;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDDLHV 10  
DB 280 SLAMDDLHV 289

## RESULT 12

ABU04358  
ID ABU04358 standard; protein; 680 AA.

AC ABU04358;

DT 29-JAN-2003 (first entry)

DE Human expressed protein tag (EPT) #1024.

XX  
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
XX protease; protease inhibitor; transporter; cytoskeletal protein;  
XX receptor; transcription factor; cancer; MHC;  
XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX Homo sapiens.

XX WO200278524-A2.

XX 10-OCT-2002.

XX 28-MAR-2002; 2002WO-US009671.

XX 28-MAR-2001; 2001US-0279495P.

XX 21-MAY-2001; 2001US-0292544P.

XX 08-AUG-2001; 2001US-0310801P.

XX 01-OCT-2001; 2001US-0326370P.

XX 04-DEC-2001; 2001US-0336780P.

XX 20-FEB-2002; 2002US-0358985P.

XX (ZYCO-) ZYCOS INC.

XX Chicx RM, Tomlinson AJ, Urban RG;

XX WPI; 2003-040607/03.

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
XX cytoskeletal proteins, receptors or transcription factors), useful for  
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
XX leukemia.

XX Example 2; SEQ ID NO 1024; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a  
XX fragment of a kinase, phosphatase, protease, protease inhibitor,  
XX transporter, cytoskeletal protein, receptor or transcription factor. The  
XX polypeptide is useful as an immunogenic composition for eliciting in a  
XX mammal an immunogenic response directed against any of the purified  
XX polypeptide. The purified polypeptide, or the antibody that binds to this

CC polypeptide, is useful for treating cancer. The polypeptide is also  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptides and  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
 CC lymphoma or leukemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: This sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pcr\_sequences

SO Sequence 680 AA;

Query Match 100.0%; Score 47; DB 6; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
 |||||  
 Db 280 SLAMDLHLV 289

# RESULT 13

ID ABG74472 standard; protein; 680 AA.

AC ABG74472;

DT 11-APR-2003 (first entry)

DE NPM/ALK fusion construct SEQ ID 4.

KM ALK; human; nucleophosmin; NPM; anaplastic lymphoma kinase; chromosome 5;  
 t(2; 5) lymphoma cell; chromosomal rearrangement; chromosome 2;  
 nucleolar phosphoprotein; centromeric; telomeric.

OS Synthetic.

Key Location/Qualifiers  
 1..117  
 /note="NPM-derived protein fragment"  
 Domain 104..115  
 /note="potential metal binding domain"  
 Protein 118..680  
 /note="ALK-derived protein fragment"  
 Domain 182..437  
 /label= ALK\_catalytic\_domain

US6451997-B1.

17-SEP-2002.

28-SEP-2000; 2000US-00670827.

03-DEC-1993; 93US-00160861.

12-OCT-1995; 95US-00542363.

19-JUN-1998; 98US-00100089.

(SUID-) ST JUDE CHILDREN'S RES HOSPITAL.

Morris SW, Look AT;

WPI; 2003-101739/09.

N-PSDB; ABG77107.  
 Kit for use in method of detecting t(2;5) chromosomal rearrangements or  
 rearrangements involving nucleolar phosphoprotein (NPM) or anaplastic  
 PT lymphoma kinase (ALK) comprises probes that hybridizes with NPM and ALK  
 PT genes.

Example 1; Fig 2A; 87bp; English.

CC This invention describes a novel kit for use in a method of detecting t(2;  
 CC 5) chromosomal rearrangements (CR) or CRs involving nucleolar  
 CC phosphoprotein (NPM) and anaplastic lymphoma kinase (ALK). The method  
 CC comprises labelled probes that hybridize to a sequence of 2442 or 6226  
 CC nucleotides or their complement, and to region of human chromosome 5/2  
 CC having wild-type NPM/ALK genes (the region is centromeric/telomeric to  
 CC nucleotide). This sequence represents a polypeptide sequence described in  
 CC the disclosure of the invention

SO Sequence 680 AA;

Query Match 100.0%; Score 47; DB 6; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
 |||||  
 Db 280 SLAMDLHLV 289

# RESULT 14

ID AAB73857 standard; protein; 1620 AA.

AC AAB73857;

DT 15-MAY-2001 (first entry)

DE Human ALK protein.

KM Human; anaplastic lymphoma kinase; ALK; nucleophosmin; NPM;  
 chromosomal translocation; cancer.

OS Homo sapiens.

US6174674-B1.

16-JAN-2001.

19-JUN-1998; 98US-00100089.

03-DEC-1993; 93US-00160861.

12-OCT-1995; 95US-00542363.

(SUID-) ST JUDE CHILDREN'S RES HOSPITAL.

Morris SW, Look AT;

WPI; 2001-243208/25.

N-PSDB; AAF76867.

Detection of chromosomal rearrangement or translocations present in t(2;  
 5) for cancer diagnosis comprises using and identifying anaplastic  
 PT lymphoma kinase and nucleophosmin genes in nucleic acid hybridization  
 PT techniques.

Claim 10; Fig 3B; 87bp; English.

CC The present sequence is given in a specification relating to a method for  
 CC detecting a chromosomal rearrangement involving a breakpoint in the  
 CC anaplastic lymphoma kinase (ALK) or nucleophosmin (NPM) gene. The method  
 CC is based on the identification and sequence determination of a novel  
 CC gene, ALK, which is fused to the gene encoding NPM in translocations  
 CC present in t(2;5) lymphoma cells. The method is useful for detecting a  
 CC t(2;5) chromosomal rearrangement, or a chromosomal rearrangement  
 CC involving the ALK gene or NPM gene. It is particularly useful in  
 CC determining if particular cells or tissues express ALK or NPM/ALK coding  
 CC sequences. The method is also useful in diagnostic assays to determine,  
 CC for example, if a mammal has cancer or a genetic predisposition to cancer  
 CC

SO Sequence 1620 AA;

Query Match 100.0%; Score 47; DB 4; Length 1620;

Best Local Similarity 100.0%; Pred. No. 3.4;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDILHY 10  
1220 SLAMDILHY 1229

## RESULT 15

AA018498 standard; protein; 1620 AA.

AA018498;

11-OCT-2002 (first entry)

Human insulin receptor signaling modifier SEQ ID NO: 18.

Human; insulin receptor signaling; insulin receptor signaling modifier;  
ISM; diabetes; metabolic syndrome; antidiabetic.

Homo sapiens.

MO200255664-A2.

18-JUL-2002.

11-JAN-2002; 2002WO-US001048.

12-JAN-2001; 2001US-0261226P.  
12-JAN-2001; 2001US-0261303P.  
12-JAN-2001; 2001US-0261304P.  
12-JAN-2001; 2001US-0261335P.  
12-JAN-2001; 2001US-0261336P.  
12-JAN-2001; 2001US-0261351P.  
12-JAN-2001; 2001US-0261456P.  
12-JAN-2001; 2001US-0261457P.  
12-JAN-2001; 2001US-0261458P.  
12-JAN-2001; 2001US-0261459P.  
12-JAN-2001; 2001US-0261461P.  
12-JAN-2001; 2001US-0261518P.  
12-JAN-2001; 2001US-0261531P.  
12-JAN-2001; 2001US-0261532P.  
12-JAN-2001; 2001US-0261589P.  
12-JAN-2001; 2001US-0261590P.  
12-JAN-2001; 2001US-0261634P.  
12-JAN-2001; 2001US-0261635P.  
12-JAN-2001; 2001US-0261697P.

(EXEL-) EXELIXIS INC.

Seidel-Dugan C, Ferguson KC, Kidd T;

WPI, 2002-599664/64.

N-PSDB; AAL48617.

Identifying an insulin receptor signaling modulator, useful as drug  
targets for treating diabetes or metabolic disorders, comprises,  
contacting an assay system comprising insulin receptor signaling  
modifiers with a test agent.

Disclosure; Page 59-66; 232pp; English.

The present invention relates to a method of identifying a candidate  
insulin receptor (INR) signaling modulating agent, involving contacting  
an assay system comprising an insulin receptor signaling modifier (ISM)  
polypeptide or nucleic acid with a test agent, and detecting a test agent  
-biased activity of the assay system. The method is useful for  
identifying candidate INR signaling modulating agents. ISM genes may be  
used as drug targets for treatment of disorders related to INR signaling  
such as diabetes or metabolic syndrome. ISM nucleic acids and  
polypeptides are useful for identifying and testing agents that modulate  
ISM function and for other applications related to the involvement of ISM

CC in INR signaling, and for identifying subjects having a predisposition to  
CC such diseases associated with INR signaling. The present sequence is an  
CC ISM protein described in the exemplification of the invention  
XX

QY Sequence 1620 AA;

Query Match 100.0%; Score 47; DB 5; Length 1620;  
Best Local Similarity 100.0%; Pred. No. 3.4;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDILHY 10  
1220 SLAMDILHY 1229

Search completed: November 30, 2004, 07:39:53  
Job time : 79.4295 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: November 30, 2004, 07:19:22 ; Search time 376.923 Seconds  
(without alignments)  
9.410 Million cell updates/sec

Title: US-10-008-377A-1  
Perfect score: 47  
Sequence: 1 SLAMDLHLV 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1575965 seqs, 354694765 residues

Total number of hits satisfying chosen parameters: 1575965

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

## Database: Published Applications-AA\*

1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*  
19: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*  
20: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	100.0	10	14	US-10-008-377-1
2	47	100.0	45	9	US-09-827-949-14
3	47	100.0	370	16	US-10-664-421-70
4	47	100.0	680	9	US-09-827-949-4
5	47	100.0	1620	9	US-09-827-949-2
6	43	91.5	9	14	US-10-008-377-3
7	39	83.0	9	14	US-10-008-377-4
8	37	78.7	391	14	US-10-369-493-2155
9	36	76.6	49	15	US-10-424-539-237107
10	36	76.6	129	17	US-10-425-115-202166
11	36	76.6	140	16	US-10-767-701-533286
12	36	76.6	253	17	US-10-425-115-297607
13	36	76.6	376	15	US-10-041-018-26

14	36	76.6	376	15	US-10-041-018-41	Sequence 41, Appl
15	36	76.6	376	15	US-10-041-018-47	Sequence 67, Appl
16	36	76.6	376	16	US-10-437-963-176494	Sequence 16494,
17	36	76.6	563	17	US-10-425-115-297605	Sequence 297605,
18	35	74.5	187	9	US-09-778-927A-68	Sequence 68, Appl
19	35	74.5	302	14	US-10-128-714-3559	Sequence 3559, Ap
20	35	74.5	334	14	US-10-128-714-8559	Sequence 8559, Ap
21	35	74.5	442	10	US-09-826-109-497	Sequence 497, Ap
22	35	74.5	442	14	US-10-020-141-6	Sequence 6, Appl
23	35	74.5	442	14	US-10-282-525-10	Sequence 10, Appl
24	35	74.5	442	14	US-10-285-567A-114	Sequence 19, Appl
25	35	74.5	442	14	US-10-372-683-49	Sequence 49, Appl
26	35	74.5	532	15	US-10-311-671-11	Sequence 11, Appl
27	35	74.5	532	16	US-10-408-765A-2102	Sequence 2102, Ap
28	34	72.3	496	15	US-10-335-977-5681	Sequence 5681, Ap
29	33	70.2	98	16	US-10-767-701-61922	Sequence 61922, A
30	33	70.2	144	15	US-10-424-539-172565	Sequence 172565,
31	33	70.2	152	15	US-10-425-114-45278	Sequence 45278, A
32	33	70.2	175	15	US-10-282-122A-45446	Sequence 45446, A
33	33	70.2	176	15	US-10-282-122A-45491	Sequence 45491, A
34	33	70.2	292	14	US-10-017-161-32	Sequence 32, Appl
35	33	70.2	292	15	US-10-343-650A-572	Sequence 572, App
36	33	70.2	323	9	US-09-886-055-163	Sequence 163, App
37	33	70.2	323	10	US-09-804-291-163	Sequence 67, Appl
38	33	70.2	323	11	US-09-981-666A-67	Sequence 78, Appl
39	33	70.2	323	14	US-10-387-629-78	Sequence 634, App
40	33	70.2	323	15	US-10-282-798-634	Sequence 314, App
41	33	70.2	323	14	US-10-343-650A-314	Sequence 37, Appl
42	33	70.2	323	17	US-10-149-826-34	Sequence 17, Appl
43	33	70.2	326	15	US-10-311-671-17	Sequence 10, Appl
44	33	70.2	328	14	US-10-010-568-10	Sequence 14, Appl
45	33	70.2	328	14	US-10-010-568-14	Sequence 14, Appl

## ALIGNMENTS

## RESULT 1

US-10-008-377-1  
Sequence 1, Application US/10008377  
Publication No. US2003015710A1  
GENERAL INFORMATION:  
APPLICANT: Gambacorti-Passerini, Carlo  
TITLE OF INVENTION: Immunogenic ALK peptides  
FILE REFERENCE: 045922/241203  
CURRENT APPLICATION NUMBER: US/10/008,377  
CURRENT FILING DATE: 2001-11-15  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-008-377-1

Query Match 100.0%; Score 47; DB 14; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.046;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 1 SLAMDLHLV 10  
DB 1 SLAMDLHLV 10  
RESULT 2  
US-09-827-949-14  
Sequence 14, Application US/09827949  
Patent No. US2001021505A1  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof

FILE REFERENCE: 0656.0400004  
CURRENT APPLICATION NUMBER: US/09/827,949  
CURRENT FILING DATE: 2001-04-09  
PRIOR APPLICATION NUMBER: US 09/670,827  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 09/100,089  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: US 08/542,363  
PRIOR FILING DATE: 1995-10-12  
PRIOR APPLICATION NUMBER: US 08/160,861  
PRIOR FILING DATE: 1993-12-03  
NUMBER OF SEQ ID NOS: 43  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 14  
LENGTH: 45  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-827-949-14

Query Match  
Best Local Similarity 100.0%; Score 47; DB 9; Length 45;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
|||||  
DB 5 SLAMDLHLV 14

RESULT 3  
US-10-664-421-70  
Sequence 70, Application US/10664421  
Publication No. US20040142864A1  
GENERAL INFORMATION:  
APPLICANT: BREMER, RYAN  
APPLICANT: IBRAHIM, PRABHA  
APPLICANT: KUMAR, ABHINAV  
APPLICANT: MANDIVAN, VALSAN V.  
APPLICANT: MILEURN, MICHAEL V.  
TITLE OF INVENTION: CRYSTAL STRUCTURE OF PIM-1 KINASE  
FILE REFERENCE: 039363/0703  
CURRENT APPLICATION NUMBER: US/10/664,421  
CURRENT FILING DATE: 2003-09-16  
PRIOR APPLICATION NUMBER: 60/412,341  
PRIOR FILING DATE: 2002-09-20  
PRIOR APPLICATION NUMBER: 60/411,398  
PRIOR FILING DATE: 2002-09-16  
NUMBER OF SEQ ID NOS: 169  
SOFTWARE: PatentIn Ver. 3.2  
SEQ ID NO 70  
LENGTH: 370  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-664-421-70

Query Match  
Best Local Similarity 100.0%; Score 47; DB 16; Length 370;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
|||||  
DB 155 SLAMDLHLV 164

RESULT 4  
US-09-827-949-4  
Sequence 4, Application US/09827949  
Publication No. US20010021505A1  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
APPLICANT: Look, A. Thomas  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof  
FILE REFERENCE: 0656.0400004  
CURRENT APPLICATION NUMBER: US/09/827,949

CURRENT FILING DATE: 2001-04-09  
PRIOR APPLICATION NUMBER: US 09/670,827  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 09/100,089  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: US 08/542,363  
PRIOR FILING DATE: 1995-10-12  
PRIOR APPLICATION NUMBER: US 08/160,861  
PRIOR FILING DATE: 1993-12-03  
NUMBER OF SEQ ID NOS: 43  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 4  
LENGTH: 680  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-827-949-4

Query Match  
Best Local Similarity 100.0%; Score 47; DB 9; Length 680;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
|||||  
DB 280 SLAMDLHLV 289

RESULT 5  
US-09-827-949-2  
Sequence 2, Application US/09827949  
Patent No. US20010021505A1  
GENERAL INFORMATION:  
APPLICANT: Look, A. Thomas  
APPLICANT: Morris, Stephan W.  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof  
FILE REFERENCE: 0656.0400004  
CURRENT APPLICATION NUMBER: US/09/827,949  
CURRENT FILING DATE: 2001-04-09  
PRIOR APPLICATION NUMBER: US 09/670,827  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 09/100,089  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: US 08/542,363  
PRIOR FILING DATE: 1995-10-12  
PRIOR APPLICATION NUMBER: US 08/160,861  
PRIOR FILING DATE: 1993-12-03  
NUMBER OF SEQ ID NOS: 43  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 2  
LENGTH: 1620  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-827-949-2

Query Match  
Best Local Similarity 100.0%; Score 47; DB 9; Length 1620;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
|||||  
DB 1220 SLAMDLHLV 1225

RESULT 6  
US-10-008-377-3  
Sequence 3, Application US/10008377  
Publication No. US20030157101A1  
GENERAL INFORMATION:  
APPLICANT: Gambacorti-Passerini, Carlo  
APPLICANT: Passoni, Lorenza  
TITLE OF INVENTION: Immunogenic ALK Peptides  
FILE REFERENCE: 045922/241203  
CURRENT APPLICATION NUMBER: US/10/008,377  
CURRENT FILING DATE: 2001-11-15

NUMBER OF SEQ ID NOS: 7  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-008-377-3

Query Match 91.5%; Score 43; DB 14; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
1 LAMLDLHV 9

RESULT 7  
US-10-008-377-4  
Sequence 4, Application US/10008377  
Publication No. US20030157101A1  
GENERAL INFORMATION:  
APPLICANT: Gambacorti-Passerini, Carlo  
APPLICANT: Passoni, Lorenza  
TITLE OF INVENTION: Immunogenic Alk Peptides  
FILE REFERENCE: 045922/241203  
CURRENT APPLICATION NUMBER: US/10/008,377  
CURRENT FILING DATE: 2001-11-15  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-008-377-4

Query Match 83.0%; Score 39; DB 14; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAMLDLHV 10  
1 LAMLDLHV 8

RESULT 8  
US-10-369-493-2155  
Sequence 2155, Application US/10369493  
Publication No. US2003023675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xiaofeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10152052/B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 2155  
LENGTH: 391  
TYPE: PRT  
ORGANISM: Schizosaccharomyces pombe  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1) (391)  
OTHER INFORMATION: unsure at all Xaa locations  
US-10-369-493-2155

Query Match 78.7%; Score 37; DB 14; Length 391;  
Best Local Similarity 88.9%; Pred. No. 1.3e+02;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
125 LAMLDLHV 133

RESULT 9  
US-10-424-599-237107  
Sequence 237107, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J  
APPLICANT: Kovacic, David K  
APPLICANT: Zhou, Yihua  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223) B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 237107  
LENGTH: 49  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_56132C.1.pcp  
US-10-424-599-237107

Query Match 76.6%; Score 36; DB 15; Length 49;  
Best Local Similarity 70.0%; Pred. No. 24;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMLDLHV 10  
20 SLAMLDLHV 29

RESULT 10  
US-10-425-115-202166  
Sequence 202166, Application US/10425115  
Publication No. US20040214272A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovacic, David K.  
APPLICANT: Zhou, Yihua  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223) B  
CURRENT APPLICATION NUMBER: US/10/425,115  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 202166  
LENGTH: 129  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: MRT4577\_115965C.1.pcp  
US-10-425-115-202166

Query Match 76.6%; Score 36; DB 17; Length 129;  
Best Local Similarity 80.0%; Pred. No. 63;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLAMLDLHV 10  
81 SLAMLDLHV 90

## RESULT 11

US-10-767-701-53329  
; Sequence 53329, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21(5353)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 53329  
; LENGTH: 140  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(140)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 13318357.pcp  
US-10-767-701-53329

## Query Match

Best Local Similarity 76.6%; Score 36; DB 16; Length 140;  
Best Local Similarity 66.7%; Pred. No. 69;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
DB 131 LAMMDLHV 139

## RESULT 12

US-10-425-115-297607  
; Sequence 297607, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(5322)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 297607  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(253)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MPT4577\_34493C.1.pcp  
US-10-425-115-297607

## Query Match

Best Local Similarity 76.6%; Score 36; DB 17; Length 253;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAMLDLHV 10  
DB 186 LAMMDLHV 194

## RESULT 13

US-10-041-018-26

Sequence 26, Application US/10041018  
; Publication No. US20040072323A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsuda, Seichi P.T.  
; APPLICANT: Hart, Elizabeth A.  
; TITLE OF INVENTION: Diterpene-Producing Unicellular Organism  
; FILE REFERENCE: P02080US1/10025547  
; CURRENT APPLICATION NUMBER: US/10/041,018  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: US 60/259880  
; NUMBER OF SEQ ID NOS: 413  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 26  
; LENGTH: 376  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-10-041-018-26

Query Match 76.6%; Score 36; DB 15; Length 376;  
Best Local Similarity 70.0%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLHV 10  
DB 250 SLAMLEIHV 259

## RESULT 14

US-10-041-018-41  
; Sequence 41, Application US/10041018  
; Publication No. US20040072323A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsuda, Seichi P.T.  
; APPLICANT: Hart, Elizabeth A.  
; TITLE OF INVENTION: Diterpene-Producing Unicellular Organism  
; FILE REFERENCE: P02080US1/10025547  
; CURRENT APPLICATION NUMBER: US/10/041,018  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: US 60/259880  
; PRIOR FILING DATE: 2001-01-05  
; NUMBER OF SEQ ID NOS: 413  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 41  
; LENGTH: 376  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-10-041-018-41

Query Match 76.6%; Score 36; DB 15; Length 376;  
Best Local Similarity 70.0%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLHV 10  
DB 250 SLAMLEIHV 259

## RESULT 15

US-10-041-018-67  
; Sequence 67, Application US/10041018  
; Publication No. US20040072323A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsuda, Seichi P.T.  
; APPLICANT: Hart, Elizabeth A.  
; TITLE OF INVENTION: Diterpene-Producing Unicellular Organism  
; FILE REFERENCE: P02080US1/10025547  
; CURRENT APPLICATION NUMBER: US/10/041,018  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: US 60/259880  
; PRIOR FILING DATE: 2001-01-05  
; NUMBER OF SEQ ID NOS: 413  
; SOFTWARE: PatentIn version 3.1  
US-10-041-018-67

Tue Nov 30 08:49:19 2004

us-10-008-377a-1.open.rapb

Page 5

; SEQ ID NO 67  
; LENGTH: 376  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-10-041-018-67

Query Match 76.6%; Score 36; DB 15; Length 376;  
Best Local Similarity 70.0%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SWAMDLHY 10  
|||: ||  
DB 250 SWAMLYTHV 259

Search completed: November 30, 2004, 08:06:34  
Job time : 377.923 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 ; Search time 25.182 Seconds

(without alignments)  
26.392 Million cell updates/sec

Title: US-10-008-377A-1

Sequence: 1 SLAMDHLHV 10

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database:

Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/prodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/ECTUS.COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	100.0	45	1	US-08-542-363-14
2	47	100.0	45	3	US-09-100-089-14
3	47	100.0	45	4	US-09-670-827-14
4	47	100.0	45	4	US-09-827-949-14
5	47	100.0	45	4	US-09-827-949-14
6	47	100.0	525	1	US-08-160-861-4
7	47	100.0	572	1	US-08-160-861-3
8	47	100.0	680	1	US-08-542-363-4
9	47	100.0	680	1	US-09-100-089-4
10	47	100.0	680	4	US-09-670-827-4
11	47	100.0	680	4	US-09-827-949-4
12	47	100.0	1620	1	US-08-542-363-2
13	47	100.0	1620	3	US-09-100-089-2
14	47	100.0	1620	4	US-09-670-827-2
15	47	100.0	1620	4	US-09-827-949-2
16	47	100.0	442	3	US-08-121-446-4
17	47	100.0	442	4	US-09-520-210-10
18	47	100.0	442	1	US-08-117-361C-1
19	47	100.0	145	4	US-09-517-999C-7009
20	47	100.0	153	4	US-09-252-991A-21109
21	47	100.0	333	4	US-08-118-270-40
22	47	100.0	333	5	PCR-US93-08528-40
23	47	100.0	385	4	US-09-543-681A-4399
24	47	100.0	385	4	US-09-489-039A-11917
25	47	100.0	385	4	US-09-543-681A-7130
26	47	100.0	391	4	US-09-270-767-43901
27	47	100.0	404	4	US-09-252-991A-18789
28	47	100.0	513	4	US-09-252-991A-30773

28	32	68.1	1047	4	US-09-529-239D-99	Sequence 99, Appl
29	31	66.0	236	4	US-08-015-985-6	Sequence 6, Appl
30	31	66.0	236	4	US-09-280-597-6	Sequence 6, Appl
31	31	66.0	254	2	US-08-685-992-9	Sequence 9, Appl
32	31	66.0	254	2	US-09-144-925-9	Sequence 9, Appl
33	31	66.0	340	4	US-09-712-368-1	Sequence 1, Appl
34	31	66.0	387	3	US-09-134-001C-3625	Sequence 3625, Ap
35	31	66.0	401	4	US-09-252-991A-19699	Sequence 19699, A
36	31	66.0	403	4	US-09-543-681A-6631	Sequence 6631, Ap
37	31	66.0	576	4	US-09-489-039A-7859	Sequence 7859, Ap
38	31	66.0	781	4	US-09-252-991A-26207	Sequence 26207, A
39	31	66.0	793	4	US-08-015-985-3	Sequence 3, Appl
40	31	66.0	793	4	US-09-280-597-3	Sequence 3, Appl
41	31	66.0	802	1	US-08-015-985-1	Sequence 1, Appl
42	31	66.0	802	4	US-09-280-597-1	Sequence 1, Appl
43	31	66.0	1032	4	US-09-252-991A-24058	Sequence 24058, A
44	31	66.0	1101	3	US-09-770-170-8	Sequence 8, Appl
45	31	66.0	2713	5	PCT-US96-01735-1	Sequence 1, Appl

#### ALIGNMENTS

RESULT 1  
US-08-542-363-14  
Sequence 14, Application US/08542363  
Patent No. 5770421

GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.

TITLE OF INVENTION: Look, A. Thomas

TITLE OF INVENTION: Ligands Thereof

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESSES:

ADDRESS: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/542,363

FILING DATE: 12-OCT-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.

REGISTRATION NUMBER: 30,353

REFERENCE/DOCKET NUMBER: 0656.0400001/SLF/GKT

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2640

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 45 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-542-363-14

Query Match: 100.0%; Score 47; DB 1; Length 45;  
Best Local Similarity: 100.0%; Pred. No. 0.096;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDHLHV 10

Db 5 SLAMDHLHV 14

RESULT 2  
US-09-100-089-14  
Sequence 14, Application US/09100089  
Patent No. 6174674  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
TITLE OF INVENTION: Ligands Thereof  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/100,089  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656,0400002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 45 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-100-089-14

Query Match 100.0%; Score 47; DB 3; Length 45;  
Best Local Similarity 100.0%; Pred. No. 0.096;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
DB 5 SLAMDLHLV 14

RESULT 3  
US-09-670-827-14  
Sequence 14, Application US/09670827  
Patent No. 6451997  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
TITLE OF INVENTION: Ligands Thereof  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC

COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/670,827  
FILING DATE: 28-Sep-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995  
APPLICATION NUMBER: US 09/100,089  
FILING DATE: 19-JUN-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656,0400003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 45 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 14:  
US-09-670-827-14

Query Match 100.0%; Score 47; DB 4; Length 45;  
Best Local Similarity 100.0%; Pred. No. 0.096;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
DB 5 SLAMDLHLV 14

RESULT 4  
US-09-827-949-14  
Sequence 14, Application US/09827949  
Patent No. 6696548  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof  
FILE REFERENCE: 0656,0400004  
CURRENT APPLICATION NUMBER: US/09/827,949  
CURRENT FILING DATE: 2001-04-09  
PRIOR APPLICATION NUMBER: US 08/670,827  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 09/100,089  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: US 08/542,363  
PRIOR FILING DATE: 1995-10-12  
PRIOR APPLICATION NUMBER: US 08/160,861  
PRIOR FILING DATE: 1993-12-03  
NUMBER OF SEQ ID NOS: 43  
SOFTWARE: Patent version 3.0  
SEQ ID NO 14  
LENGTH: 45  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-827-949-14

Query Match 100.0%; Score 47; DB 4; Length 45;  
Best Local Similarity 100.0%; Pred. No. 0.096;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLAMDLHLV 10  
|||||||  
DB 5 SLAMDLHLV 14

RESULT 5  
US-08-160-861-4

Sequence 4, Application US/08160861  
Patent No. 5529295

GENERAL INFORMATION:

APPLICANT: MORRIS, STEPHAN W

TITLE OF INVENTION: NOVEL FUSION NUCLEIC ACID SEQUENCES AND

TITLE OF INVENTION: FUSION PROTEINS PRESENT IN HUMAN t(2:5) LYMPHOMA, METHODS

TITLE OF INVENTION: OF DETECTION AND USES THEREOF

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX

STREET: 1100 NEW YORK AVE NW SUITE 600

CITY: WASHINGTON

STATE: D.C.

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/160,861

FILING DATE: 02-DEC-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: MILLMAN, ROBERT A

REGISTRATION NUMBER: 36217

REFERENCE/DOCKET NUMBER: 0656.0400000

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2678

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 525 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-160-861-4

Query Match 100.0%; Score 47; DB 1; Length 525;

Best Local Similarity 100.0%; Pred. No. 1.2;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLAMDLHLV 10  
|||||||

DB 280 SLAMDLHLV 289

RESULT 6  
US-08-160-861-3

Sequence 3, Application US/08160861

Patent No. 5529295

GENERAL INFORMATION:

APPLICANT: MORRIS, STEPHAN W

TITLE OF INVENTION: NOVEL FUSION NUCLEIC ACID SEQUENCES AND

TITLE OF INVENTION: FUSION PROTEINS PRESENT IN HUMAN t(2:5) LYMPHOMA, METHODS

TITLE OF INVENTION: OF DETECTION AND USES THEREOF

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SLAMDLHLV 10  
|||  
Db 280 SLAMDLHLV 289

RESULT 8  
US-09-100-089-4  
; Sequence 4, Application US/09100089  
; Patent No. 6174674  
; GENERAL INFORMATION:  
; APPLICANT: Morris, Stephan W.  
; APPLICANT: Look, A. Thomas  
; TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
; TITLE OF INVENTION: Ligands Thereof  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 New York Avenue, N.W., Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/100,089  
; FILING DATE: Herewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/160,861  
; FILING DATE: 03-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/542,363  
; FILING DATE: 12-OCT-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fox, Samuel L.  
; REGISTRATION NUMBER: 30,353  
; REFERENCE/DOCKET NUMBER: 0656.0400002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-2500  
; TELEFAX: 202-371-2540  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 680 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-100-089-4

Query Match 100.0%; Score 47; DB 3; Length 680;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
|||  
Db 280 SLAMDLHLV 289

RESULT 9  
US-09-670-827-4  
; Sequence 4, Application US/09670827  
; Patent No. 6451997  
; GENERAL INFORMATION:  
; APPLICANT: Morris, Stephan W.  
; APPLICANT: Look, A. Thomas  
; TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
; TITLE OF INVENTION: Ligands Thereof  
; NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/670,827  
FILING DATE: 28-SEP-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995  
APPLICATION NUMBER: US 09/100,089  
FILING DATE: 19-JUN-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2500  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 680 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-670-827-4

Query Match 100.0%; Score 47; DB 4; Length 680;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
|||  
Db 280 SLAMDLHLV 289

RESULT 10  
US-09-827-949-4  
; Sequence 4, Application US/09827949  
; Patent No. 6696548  
; GENERAL INFORMATION:  
; APPLICANT: Morris, Stephan W.  
; APPLICANT: Look, A. Thomas  
; TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof  
; FILE REFERENCE: 0656.0400004  
; CURRENT APPLICATION NUMBER: US/09/827,949  
; PRIOR FILING DATE: 2001-04-09  
; PRIOR APPLICATION NUMBER: US 09/670,827  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 09/100,089  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: US 08/542,363  
; PRIOR FILING DATE: 1995-10-12  
; PRIOR APPLICATION NUMBER: US 08/160,861  
; PRIOR FILING DATE: 1993-12-03  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn Version 3.0  
; SEQ ID NO: 4  
; LENGTH: 680  
; TYPE: PRT  
; ORGANISM: Homo sapiens

Tue Nov 30 08:49:18 2004

us-10-008-377a-1.open.rat

Page 5

US-09-827-949-4

Query Match 100.0%; Score 47; DB 4; Length 680;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
DB 280 SLAMDLHLV 289

RESULT 11  
US-08-542-363-2  
Sequence 2, Application US/08542363  
Patent No. 5770421  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
APPLICANT: Look, A. Thomas  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
TITLE OF INVENTION: Ligands Thereof  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/542,363  
FILING DATE: 12-OCT-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400001/SLF/SKT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2540  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ. ID NO.: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1620 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-542-363-2  
Query Match 100.0%; Score 47; DB 1; Length 1620;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
DB 1220 SLAMDLHLV 1229

RESULT 12  
US-09-100-089-2  
Sequence 2, Application US/09100089  
Patent No. 6174674  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
APPLICANT: Look, A. Thomas  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
TITLE OF INVENTION: Ligands Thereof  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:

ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/100,089  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ. ID NO.: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1620 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-100-089-2

Query Match 100.0%; Score 47; DB 3; Length 1620;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
DB 1220 SLAMDLHLV 1229

RESULT 13  
US-09-670-827-2  
Sequence 2, Application US/09670827  
Patent No. 6451997  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
APPLICANT: Look, A. Thomas  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
TITLE OF INVENTION: Ligands Thereof  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/670,827  
FILING DATE: 28-Sep-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995  
APPLICATION NUMBER: US 09/100,089  
FILING DATE: 19-JUN-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1620 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-670-827-2

Query Match 100.0%; Score 47; DB 4; Length 1620;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
DB 1220 SLAMDLHLV 1229

RESULT 14  
US-09-827-949-2  
Sequence 2, Application US/09827949  
Patent No. 6596549  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
APPLICANT: Look, A. Thomas  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof  
FILE REFERENCE: 0656.0400004  
CURRENT APPLICATION NUMBER: US/09/827,949  
CURRENT FILING DATE: 2001-04-09  
PRIOR APPLICATION NUMBER: US 09/670,827  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 09/100,089  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: US 08/542,363  
PRIOR FILING DATE: 1995-10-12  
PRIOR APPLICATION NUMBER: US 08/160,861  
PRIOR FILING DATE: 1993-12-03  
NUMBER OF SEQ ID NOS: 43  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO: 2  
LENGTH: 1620  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-827-949-2

Query Match 100.0%; Score 47; DB 4; Length 1620;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
DB 1220 SLAMDLHLV 1229

RESULT 15  
US-08-121-446-4  
Sequence 4, Application US/08121446  
Patent No. 6313276  
GENERAL INFORMATION:

APPLICANT: IMURA, HIROO  
APPLICANT: NAKANO, KAZUWA  
APPLICANT: NAKANISHI, SHIGERU  
TITLE OF INVENTION: A HUMAN ENDOTHELIN RECEPTOR  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 Page Mill Road  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/121,446  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/911,684  
FILING DATE: 10-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: CIOTTI, THOMAS E.  
REGISTRATION NUMBER: 21,013  
REFERENCE/DOCKET NUMBER: 29900-20324.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 442 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-121-446-4

Query Match 74.5%; Score 35; DB 3; Length 442;  
Best Local Similarity 70.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMDLHLV 10  
DB 142 SLAMDLHLV 151

Search completed: November 30, 2004, 07:19:15  
Job time : 26.1282 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 ; Search time 20 Seconds

(without alignments)  
48.108 Million cell updates/sec

Title: US-10-008-377A-2

Perfect score: 52

Sequence: 1 GVLMWEIFSL 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

283416

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: p1r1.\*  
2: p1r2.\*  
3: p1r3.\*  
4: p1r4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	100.0	367	2 B56598	endothelial kinase
2	52	100.0	464	2 A36625	protein-tyrosine k
3	52	100.0	576	2 S12792	protein-tyrosine k
4	52	100.0	680	4 I38491	nucleophosmin/ana
5	52	100.0	864	2 A48266	protein-tyrosine k
6	52	100.0	888	2 I58378	protein kinase -
7	52	100.0	1088	1 A48993	platelet-derived g
8	52	100.0	1298	2 A48993	protein-tyrosine k
9	52	100.0	1330	2 A49010	embryonic receptor
10	52	100.0	1333	2 I78875	receptor tyrosine
11	52	100.0	1336	2 I60598	Fit-1 tyrosine kin
12	52	100.0	1338	2 G09982	protein-tyrosine k
13	52	100.0	1348	2 S16566	vascular endotheli
14	52	100.0	1356	2 J01402	protein-tyrosine k
15	52	100.0	1363	2 I58375	protein-tyrosine k
16	52	100.0	1367	2 A41228	protein-tyrosine k
17	52	100.0	1379	2 J04954	vascular endotheli
18	52	100.0	1621	2 T10200	protein-tyrosine k
19	51	98.1	1443	2 B39061	protein-tyrosine k
20	51	98.1	160	2 A39061	protein-tyrosine k
21	51	98.1	457	2 A42659	platelet-derived g
22	51	98.1	941	1 TYWVMD	protein-tyrosine k
23	51	98.1	954	2 I51703	C-kit-related kina
24	51	98.1	972	1 TYWVMD	macrophage colony-
25	51	98.1	975	2 T10816	macrophage colony-
26	51	98.1	976	2 TYWVMD	macrophage colony-
27	51	98.1	978	2 S16385	macrophage colony-
28	51	98.1	980	1 TVCTMD	macrophage colony-
29	51	98.1	992	2 A39931	protein-tyrosine k

30	51	98.1	993	2 A36873	protein-tyrosine k
31	51	98.1	1000	2 S18827	Plt3 protein - mou
32	51	98.1	1087	2 I51552	platelet-derived g
33	51	98.1	1089	1 PFHUGA	platelet-derived g
34	51	98.1	1089	1 S37127	platelet-derived g
35	50	96.2	478	2 T32476	probable protein-t
36	49	94.2	388	2 I51023	fibroblast growth
37	49	94.2	729	2 A56795	fibroblast growth
38	49	94.2	733	2 I49293	fibroblast growth
39	49	94.2	737	2 S38579	fibroblast growth
40	49	94.2	800	1 TVHJ2F	fibroblast growth
41	49	94.2	800	2 A48991	heparin-binding gr
42	49	94.2	806	1 TVHJ2F	fibroblast growth
43	49	94.2	806	2 A35963	protein-tyrosine k
44	49	94.2	812	1 A36477	fibroblast growth
45	49	94.2	814	1 A39752	fibroblast growth

#### ALIGNMENTS

##### RESULT 1

B56598  
endothelial kinase Quek2 - quail (fragment)  
NAlternate names: Vascular endothelial growth factor receptor homolog Quek2  
CSpecies: Coturnix coturnix (quail)  
CDate: 11-Aug-1995 #sequence\_revision 11-Aug-1995 #ext\_change 16-Aug-2004  
CAccession: B56598  
RReichmann, A.; Marcelle, C.; Breant, C.; Le Douarin, N.M.  
Mech. Dev. 42, 33-48, 1993  
ATitle: Two molecules related to the VEGF receptor are expressed in early endothelial  
AReference number: A56598; WUID:93378866; PMID:8396413  
AContents: sep. japonica, E4 embryo  
AAccession: B56598  
AStatus: Preliminary  
AMolecule type: mRNA  
AResidues: 1-367 <EIC>  
AReferences: UNIPROT:Q91356; GB:S65207; NID:G410682; PTDN:AA828128.1; PTD:G41068  
ANote: sequence extracted from NCBI backbone (NCBIN:137164, NCBI:137165)  
CKeywords: Arp; growth factor receptor  
CKeywords: Arp; growth factor receptor  
P.1-176/Domain: protein kinase homology (fragment) <KIN>

Query Match Score 52; DB 2; Length 367;

Best local similarity 100.0%; Pred. No. 0.13;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10  
|||||||  
DB 101 GVLMWEIFSL 110

##### RESULT 2

A36625  
protein-tyrosine kinase (BC 2.7.1.112) Itk (version 2) - human  
CSpecies: Homo sapiens (man)  
CDate: 11-Feb-1993 #sequence\_revision 11-Feb-1993 #ext\_change 04-Feb-2000  
CAccession: A36625  
RMaru, Y.; Hirai, H.; Takaku, F.  
Oncogene Res. 5, 199-204, 1990  
ATitle: Human Itk: gene structure and preferential expression in human leukemic cells.  
AReference number: A36625; WUID:90206632; PMID:2320375  
AAccession: A36625  
AStatus: Preliminary  
AMolecule type: mRNA  
AResidues: 1-464 <MAR>  
AReferences: GB:X52213; NID:G34421; PTDN:CA364460.1; PTD:G34422  
CKeywords: Arp; phosphotyrosinase; tyrosine-specific protein kinase  
P.109-384/Domain: protein kinase homology <KIN>  
P.117-125/Region: protein kinase ATP-binding motif

Query Match Score 52; DB 2; Length 464;

Best Local Similarity 100.0%; Pred. No. 0.17;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10  
DB 310 GVLMWEIFSL 319

## RESULT 3

S12792  
protein-tyrosine kinase (EC 2.7.1.112) ltk - mouse

N/Alternate names: leukocyte tyrosine kinase

C/Species: Mus musculus (house mouse)

C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-2004

C/Accession: S12792; S00904

R/Bernards, A.; de la Monte, S.M.

EMBO J. 9, 2279-2287, 1990

A/Title: The ltk receptor tyrosine kinase is expressed in pre-B lymphocytes and cerebral

A/Reference number: S12792; PMID:90291994; PMID:2357970

A/Accession: S12792

A/Molecule type: mRNA

A/Residues: 1-576 <DB>

A/Cross-references: UNIPROT:P08923; EMBL:X52621; NID:G52948; PIDN:CAA36848.1; PID:G43792

R/Bernards, A.; Bauskin, A.R.

Nature 333, 672-676, 1988

A/Title: Leukocytes express a novel gene encoding a putative transmembrane protein-kinase

A/Reference number: S00904; PMID:88332962; PMID:2836739

A/Accession: S00904

A/Molecule type: mRNA

A/Residues: 101-476, 'V', 478-562, 'H', 564-576 <BEN>

A/Cross-references: EMBL:X07984

C/Genetics:

A/Start codon: CUG

C/Superfamily: protein kinase homology

C/Keywords: ATP, autophosphorylation, glycoprotein, membrane protein, phosphoprotein, p

F/192-468/Domain: protein kinase homology <KIN>

F/200-208/Region: protein kinase ATP-binding motif

## Query Match

Best Local Similarity 100.0%; Score 52; DB 2; Length 576;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## RESULT 4

138491  
nucleophosmin/anaplastic lymphoma kinase mutant fusion protein - human

C/Species: Homo sapiens (man)

C/Date: 15-Feb-1996 #sequence\_revision 15-Feb-1996 #text\_change 20-Apr-2000

C/Accession: 138491

R/Korits, S.W.; Kiststein, M.N.; Valentine, M.B.; Dittmer, K.G.; Shapiro, D.N.; Salzman,

Science 263, 1281-1284, 1994

A/Title: Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin'

A/Reference number: A53182; PMID:94167588; PMID:8122112

A/Accession: 138491

A/Molecule type: mRNA

A/Residues: 1-560 <NOR>

A/Cross-references: EMBL:U04946; NID:G609341; PIDN:AAA8688.1; PID:G609342

C/Comment: This sequence is the chimeric product of a translocation mutation.

C/Genetics:

A/Gene: NPM1/ALK

A/Map position: 5/2p23-2p23

C/Keywords: fusion protein

Query Match 100.0%; Score 52; DB 4; Length 680;

Best Local Similarity 100.0%; Pred. No. 0.25; 0; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 0;

DB 376 GVLMWEIFSL 385

## RESULT 5

A48266  
protein-tyrosine kinase (EC 2.7.1.112) ltk - human

N/Alternate names: protein-tyrosine kinase tyk1

C/Species: Homo sapiens (man)

C/Date: 16-Feb-1994 #sequence\_revision 12-Apr-1996 #text\_change 09-Jul-2004

C/Accession: A48266; S17452; A60189

R/Toyoshima, H.; Kozutsumi, H.; Maru, Y.; Hagiwara, K.; Futuya, A.; Mich, H.; Hanai, N.;

Proc. Natl. Acad. Sci. U.S.A. 90, 3404-3408, 1993

A/Title: Differently spliced cDNAs of human leukocyte tyrosine kinase receptor tyrosine

A/Reference number: A48266; PMID:93296146; PMID:7685902

A/Accession: A48266

A/Molecule type: mRNA

A/Residues: 1-864 <TOY>

A/Cross-references: UNIPROT:P29376; GB:D16105; NID:G440854; PIDN:BAAC3679.1; PID:dl00419

A/Experimental source: placenta

A/Note: sequence modified after extraction from NCBI backbone

A/Note: the authors translated the codon CAG for residue 42 as Arg

R/Krolewski, J.J.; Dalla-Favera, R.

EMBO J. 10, 2911-2919, 1991

A/Title: The ltk gene encodes a novel receptor-type protein tyrosine kinase.

A/Reference number: S17452; PMID:92007725; PMID:1655406

A/Accession: S17452

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-41, 'R', 43-219, 'L', 221-273, 335-864 <KRO>

A/Cross-references: EMBL:X60702; NID:G34419; PIDN:CAA43113.1; PID:G34420

R/Krolewski, J.J.; Lee, R.; Eddy, R.; Shows, T.B.; Dalla-Favera, R.

Oncogene 5, 277-282, 1990

A/Title: Identification and chromosomal mapping of new human tyrosine kinase genes.

A/Reference number: A60189; PMID:90191712; PMID:2156206

A/Accession: A60189

A/Status: preliminary; not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 608-716 <KR2>

C/Genetics:

A/Map position: 15q15.1-15q15.2

C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

C/Keywords: alternative splicing, ATP, phosphotransferase, tyrosine-specific protein kin

F/508-784/Domain: protein kinase homology <KIN>

F/516-524/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 52; DB 2; Length 864;

Best Local Similarity 100.0%; Pred. No. 0.32; 0; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 0;

## RESULT 6

158378  
tyrosine kinase - mouse

C/Species: Mus musculus (house mouse)

C/Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 16-Aug-2004

C/Accession: 158378

R/Snijders, A.J.; Haase, V.H.; Bernards, A.

Oncogene 8, 27-35, 1993

A/Title: Four tissue-specific mouse ltk mRNAs predict tyrosine kinases that differ upstr

A/Reference number: 158378; PMID:93141274; PMID:6380920

A/Accession: 158378

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-888 <RNS>

A/Cross-references: UNIPROT:P08923; GB:M90470; NID:G198885; PIDN:AAA9451.1; PID:G19888

C:Genetics:  
A:Gene: Itk  
C:Superfamily: protein kinase homology  
C:Keywords: ATP  
F:504-780/Domain: protein kinase homology <KIN>  
F:512-520/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 52; DB 2; Length 888;  
Best Local Similarity 100.0%; Pred. No. 0.32;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMEIFSL 10  
DB 706 GVLMEIFSL 715

# RESULT 7

## PERITGA

platelet-derived growth factor receptor alpha precursor - rat  
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text\_change 09-Jul-2004

C:Accession: A34720; S33767; S25100

R:Lee, K.H.; Bowen-Pope, D.P.; Reed, R.R.

Mol. Cell. Biol. 10, 2237-2246, 1990

A:Title: Isolation and characterization of the alpha platelet-derived growth factor rece

A:Reference number: A34710; MUID:90220609; PMID:2157969

A:Accession: A34710

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-1088 <LE>

A:Cross-references: UNIPROT:P20786; GB:M33837; NID:G2029299; PIDN:AAA40743.1; PID:G202930

A>Note: in the authors' translation an additional residue, Val, is shown after position

R:Herren, B.; Weyer, K.A.; Ronge, M.; Loetscher, P.; Pech, M.

Biochim. Biophys. Acta 1173, 294-302, 1993

A:Title: Conservation in sequence and affinity of human and rodent PDGF ligands and rece

A:Reference number: S33764; MUID:93305723; PMID:8318539

A:Accession: S33767

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 33-149; 'R', 151-518; 'T', 520-523 <HER1>

A:Cross-references: EMBL:Z14118; NID:G56863; PIDN:CAA78488.1; PID:G56864

A:Experimental source: strain Sprague Dawley

R:Herren, B.; Weyer, K.A.; Ronge, M.; Loetscher, P.; Pech, M.

submitted to the EMBL Data Library, July 1992

A:Description: Cross-species conservation in sequence and function of PDGF ligands and r

A:Reference number: S25096

A:Accession: S25100

A:Molecule type: mRNA

A:Residues: 33-149; 'R', 151-518; 'T', 520-523 <HER2>

A:Cross-references: EMBL:Z14118; NID:G56863; PIDN:CAA78488.1; PID:G56864

A:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;

C:Keywords: ATP; autophosphorylation; glycoprotein; heterodimer; homodimer; phosphoprote

F:1-33/Domain: signal sequence #status predicted <SIG>

F:24-108/Region: platelet-derived growth factor receptor alpha #status predicted <MAT>

F:24-523/Domain: extracellular #status predicted <EX1>

F:41-101/Domain: immunoglobulin homology <IMM1>

F:142-190/Domain: immunoglobulin homology <IMM2>

F:227-291/Domain: immunoglobulin homology <IMM3>

F:427-502/Domain: immunoglobulin homology <IMM4>

F:524-547/Domain: transmembrane #status predicted <TMN>

F:548-1088/Domain: intracellular #status predicted <INT>

F:590-955/Domain: protein kinase ATP-binding motif

F:148-991/Region: protein kinase ATP-binding motif

F:175/6,88,102,178,352,358,457,467/Binding site: carbohydrate (Asn) (covalent) #status p

F:626/Active site: Lys #status predicted

F:848/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 100.0%; Score 52; DB 1; Length 1088;  
Best Local Similarity 100.0%; Pred. No. 0.4;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMEIFSL 10  
DB 881 GVLMEIFSL 890

# RESULT 8

A48999 protein-tyrosine kinase (EC 2.7.1.112) Flt4 precursor - human

N:Alternate names: class III receptor tyrosine kinase FLT4; fms-like tyrosine kinase 4

C:Species: Homo sapiens (man)

C:Date: 21-Jan-1994 #sequence revision 18-Nov-1994 #text\_change 09-Jul-2004

C:Accession: A48999; A44930; G02316; S36130; A42010

R:Paasola, K.; Aprelikova, O.; Korhonen, U.; Kaipainen, A.; Pertovaara, L.; Alitalo, R.

Cancer Res. 52, 5738-5743, 1992

A:Title: FLT4 receptor tyrosine kinase contains seven immunoglobulin-like loops and is

A:Reference number: A48999; MUID:93007958; PMID:1327515

A:Accession: A48999

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-1298 <PAJ>

A:Cross-references: UNIPROT:P35916; PIDN:AA823636.1; PID:G257352

A:Experimental source: HEL erythroleukemia cells

R:Aprelikova, O.; Paasola, K.; Partanen, J.; Armstrong, E.; Alitalo, R.; Bailey, S.K.;

Cancer Res. 52, 746-748, 1992

A:Title: FLT4, a novel class III receptor tyrosine kinase in chromosome 5q31-qter.

A:Reference number: A44930; MUID:92119639; PMID:1310071

A:Accession: A44930

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 761-1190 <APR>

A:Cross-references: GB:X68203; NID:G31433

A>Note: sequence extracted from NCBI backbone (NCBIP:78155)

R:Wood, W.I.

submitted to the EMBL Data Library, December 1995

A:Reference number: H01039

A:Accession: G02316

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1298 <MOO>

A:Cross-references: EMBL:U43143; NID:G1150990; PIDN:AA85215.1; PID:G1150991

R:Galland, F.

submitted to the EMBL Data Library, December 1992

A:Reference number: S36130

A:Accession: S36130

A:Molecule type: mRNA

A:Residues: 1-23; 'D', 25-744; 'P', 746-751; 'RR', 754-869; 'Q', 891-1127; 'V', 1129-1145; 'H', 1147

A:Cross-references: EMBL:X69878; NID:G297049; PIDN:CAA49505.1; PID:G297050

R:Galland, F.; Karamysheva, A.; Mattei, M.G.; Rosnet, O.; Marchetto, S.; Birnbaum, D.

Genomics 13, 475-478, 1992

A:Title: Chromosomal localization of FLT4, a novel receptor-tyrosine kinase gene.

A:Reference number: A42010; MUID:92307693; PMID:1319394

A:Accession: A42010

A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tr

A:Molecule type: mRNA

A:Residues: 776-869; 'Q', 891-1127; 'V', 1129-1145; 'H', 1147-1163; 'D', 1165-1200 <GAL2>

C:Genetics:

A:Gene: GDB:FLT4

A:Cross-references: GDB:128732; OMIM:136352

A:Map position: 5q34-5q35

A:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

C:Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; transmembrane

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-1298/Product: protein-tyrosine kinase FLT4 #status predicted <MAT>

F:843-1176/Domain: protein kinase homology <KIN>

F:851-859/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 52; DB 2; Length 1298;  
Best Local Similarity 100.0%; Pred. No. 0.48;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMEIFSL 10  
DB 1 GVLMEIFSL 10

Db 1101 GVLMEIFSL 1110

## RESULT 9

S49010

embryonic receptor kinase - mouse

C:Species: Mus musculus (house mouse)

C&gt;Date: 07-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 16-Aug-2004

C:Accession: S49010

R:Choi, K.; Wall, C.; Hanraty, R.; Keller, G.

Oncogene 9, 1261-1266, 1994

A:Title: Isolation of a gene encoding a novel receptor tyrosine kinase from differentiat

A:Reference number: S49010; MUID:94181281; PMID:8134130

A:Accession: S49010

A:Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-1330 &lt;CHO&gt;

A:Cross-references: UNIPROT:P35969; EMBL:X78568; NID:G510664; PIDN:CAAS5311.1; PID:G5106

C:Superfamily: protein kinase homology

C:Keywords: ATP

F:831-839/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 52; DB 2; Length 1330;

Best Local Similarity 100.0%; Pred. No. 0.49; Mismatches 0; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMEIFSL 10

Db 1083 GVLMEIFSL 1092

## RESULT 10

178875

receptor tyrosine kinase - mouse

C:Species: Mus musculus (house mouse)

C&gt;Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 16-Aug-2004

C:Accession: 178875

R:Finamore, H.; Kelleher, K.; Morris, G.E.; Bean, K.; Merberg, D.M.; Kriz, R.; Morris, J

Oncogene 8, 2293-2298, 1993

A:Title: Molecular cloning of murine FLT and FLT4

A:Reference number: 158375; MUID:93330572; PMID:8393164

A:Accession: 178875

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1333 &lt;RES&gt;

A:Cross-references: UNIPROT:P35969; GB:L07297; NID:G293782; PIDN:AAA40078.1; PID:G293783

C:Genetics:

A:Gene: FLT

C:Superfamily: protein kinase homology

C:Keywords: ATP

F:826-1161/Domain: protein kinase homology &lt;KIN&gt;

F:834-842/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 52; DB 2; Length 1333;

Best Local Similarity 100.0%; Pred. No. 0.49; Mismatches 0; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMEIFSL 10

Db 1086 GVLMEIFSL 1095

## RESULT 11

160599

F1t-1 tyrosine kinase receptor - rat

C:Species: Rattus norvegicus (Norway rat)

C&gt;Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 16-Aug-2004

C:Accession: 160598

R:Yamane, A.; Seetharam, L.; Yamaguchi, S.; Gotch, N.; Takahashi, T.; Neufeld, G.; Shibu

Oncogene 9, 2683-2690, 1994

A:Title: A new communication system between hepatocytes and sinusoidal endothelial cells

A:Reference number: 160598; MUID:9433623; PMID:8056332

A:Accession: 160598

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1336 &lt;RES&gt;

A:Cross-references: UNIPROT:P53767; GB:D28498; NID:G511662; PIDN:BAAS5857.1; PID:G600379

C:Superfamily: protein kinase homology

C:Keywords: ATP

F:825-1161/Domain: protein kinase homology &lt;KIN&gt;

F:833-841/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 53; DB 2; Length 1336;

Best Local Similarity 100.0%; Pred. No. 0.49; Mismatches 0; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMEIFSL 10

Db 1086 GVLMEIFSL 1095

## RESULT 12

S09982

protein-tyrosine kinase (EC 2.7.1.112) flt1 precursor - human

N:Alternate names: receptor-type tyrosine kinase flt

C:Species: Homo sapiens (man)

C&gt;Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 16-Aug-2004

C:Accession: S09982

R:Shibuya, M.; Yamaguchi, S.; Yamane, A.; Ikeda, T.; Tojo, A.; Matsushima, H.; Sato, M.

Oncogene 5, 519-524, 1990

A:Title: Nucleotide sequence and expression of a novel human receptor-type tyrosine kina

A:Reference number: S09982; MUID:90221591; PMID:2158038

A:Accession: S09982

A:Molecule type: mRNA

A:Residues: 1-1338 &lt;SH1&gt;

A:Cross-references: UNIPROT:P17948; EMBL:X51602; NID:G31431; PIDN:CAAS5946.1; PID:G31432

C:Genetics:

A:Gene: GDB:FLTI

A:Cross-references: GDB:120616; OMIM:165070

A:Map position: 13q12-13q12

C:Superfamily: protein kinase homology

C:Keywords: ATP; autophosphorylation; glycoprotein; membrane protein; phosphoprotein; ph

F:1-32/Domain: signal sequence #status predicted &lt;Sig&gt;

F:23-1338/Product: protein-tyrosine kinase flt #status predicted &lt;Mat&gt;

F:825-1161/Domain: protein kinase homology &lt;KIN&gt;

F:833-841/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 52; DB 2; Length 1338;

Best Local Similarity 100.0%; Pred. No. 0.49; Mismatches 0; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMEIFSL 10

Db 1086 GVLMEIFSL 1095

## RESULT 13

S51656

vascular endothelial growth factor receptor 1 precursor - Japanese quail

N:Alternate names: quail endothelial kinase 1 (Quek 1); vascular endothelial growth fact

C:Species: Coturnix coturnix japonica (Japanese quail)

C&gt;Date: 07-May-1995 #sequence\_revision 01-Sep-1995 #text\_change 16-Aug-2004

C:Accession: J04953; A56598; I51162; S51656

R:Richmann, A.; Marcelle, C.; Breant, C.; Le Douarin, N.M.

Gene 174, 3-8, 1996

A:Title: Molecular cloning of Quek 1 and 2, two quail vascular endothelial growth factor

A:Reference number: J04953; MUID:97017121; PMID:8863722

A:Accession: J04953

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-1348 &lt;ETC1&gt;

A:Note: submitted to the EMBL Data Library, December 1994

Mech. Dev. 42, 33-48, 1993

A>Title: Two molecules related to the VEGF receptor are expressed in early endothelial cells  
 A:Reference number: A56598; MUID:93378866; PMID:8396413  
 A:Contents: E16 spinal cord  
 A:Accession: A56598  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 910-1348 <E1C2>  
 A:CROSS-references: GB:565205; NID:G410680; PIND:AA828127.1; PID:G410681  
 A>Note: sequence extracted from NCBI backbone (NCBITN:137162, NCBIPI:137163)  
 R:Marcelle, C.; Eichmann, A.  
 Oncogene 7, 2479-2487, 1992  
 A>Title: Molecular cloning of a family of protein kinase genes expressed in the avian embryo  
 A:Reference number: 150595; MUID:93096482; PMID:1281306  
 A:Accession: 151162  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1023-1079 <MAR>  
 A:CROSS-references: EMBL:X69694; NID:G395226; PIND:CAA49364.1; PID:G938278  
 A>Note: the species is not identified by the authors; the most probable species is shown  
 C:Comment: This protein is an endothelial-specific receptor and binds vascular endothelial  
 C:Superfamily: protein kinase homology  
 C:Keywords: ATP; embryo; growth factor receptor; transmembrane protein  
 F:1-19/Domain: signal sequence; status predicted <SIG>  
 F:20-1348/Product: vascular endothelial growth factor receptor 1 #status predicted <MAT>  
 F:1756-777/Domain: transmembrane #status predicted <TM>  
 F:823-1160/Domain: protein kinase homology <KIN>  
 F:831-839/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 52; DB 2; Length 1348;  
 Best Local Similarity 100.0%; Pred. No. 0.49; Mismatches 0; Indels 0; Gaps 0;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMEIFSL 10  
 |||||  
 DB 1085 GVLMEIFSL 1094

## RESULT 14

protein-tyrosine kinase (EC 2.7.1.112) KDR - human  
 C:Species: Homo sapiens (man)  
 C>Date: 30-Sep-1993 #sequence\_revision 28-Aug-1998 #text\_change 16-Aug-2004  
 C:Accession: JCI1402; 158357  
 R:Teram, B. I.; Dougher-Vernaz, M.; Carrion, M. E.; Dimetrov, D.; Armellino, D. C.; Gosh  
 Biochem. Biophys. Res. Commun. 187, 1579-1586, 1992  
 A>Title: Identification of the KDR tyrosine kinase as a receptor for vascular endothelial  
 A:Reference number: JCI1402; MUID:93038635; PMID:1147831  
 A:Accession: JCI1402  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-48, 'I', 50-616, 'E', 618-700, 'E', 702-748, 'F', 750-1263, 'I', 1265-1290, 'P', 1292-  
 A:CROSS-references: UNIPROT:P35968; EMBL:X61656; NID:G31717  
 R:Teram, B. I.; Carrion, M. E.; Kovacs, E.; Rasmussen, B. A.; Eddy, R. L.; Shows, T. B.  
 Oncogene 6, 1677-1683, 1991  
 A>Title: Identification of a new endothelial cell growth factor receptor tyrosine kinase  
 A:Reference number: 158357; MUID:92018939; PMID:1656371  
 A:Accession: 158357  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 3-1356 <TER2>  
 A:CROSS-references: GB:104947; NID:G186674; PIND:AA59459.1; PID:G186675  
 C:Comment: This enzyme acts as a receptor for vascular endothelial cell growth factor.  
 C:Genetics:  
 A:Gene: GDB:KDR  
 A:CROSS-references: GDB:127921; OMIM:191306  
 A:Map position: 4q12-4q12  
 C:Superfamily: protein kinase homology  
 C:Keywords: ATP; phosphotransferase; transmembrane protein; tyrosine-specific protein ki  
 F:166-790/Domain: transmembrane #status predicted <TM>  
 F:832-1167/Domain: protein kinase homology <KIN>  
 F:840-848/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 52; DB 2; Length 1356;

Best Local Similarity 100.0%; Pred. No. 0.5;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMEIFSL 10  
 |||||  
 DB 1092 GVLMEIFSL 1101

## RESULT 15

protein-tyrosine kinase (EC 2.7.1.112) FLT4 - mouse  
 C:Species: Mus musculus (house mouse)  
 C>Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 16-Aug-2004  
 C:Accession: 158375; B42010  
 R:Finerty, H.; Kellieher, K.; Morris, G. E.; Bean, K.; Metzberg, D. M.; Kitz, R.; Morris,  
 Oncogene 8, 2293-2298, 1993  
 A>Title: Molecular cloning of murine FLT and FLT4  
 A:Reference number: 158375; MUID:93305072; PMID:8393164  
 A:Accession: 158375  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1363 <RES>  
 A:CROSS-references: UNIPROT:P35917; GB:107296; NID:G293780; PIND:AA40077.1; PID:G29378  
 R:Galani, F.; Karamysheva, A.; Nattei, M. G.; Rosnet, O.; Marchetto, S.; Birnbaum, D.  
 Genomics 13, 475-478, 1992  
 A>Title: Chromosomal localization of FLT4, a novel receptor-type tyrosine kinase gene.  
 A:Reference number: A42010; MUID:92307693; PMID:1319394  
 A:Accession: B42010  
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tr  
 A:Molecule type: DNA  
 A:Residues: 1033-1072 <GNL>  
 C:Genetics:  
 A:Gene: FLT4  
 C:Superfamily: protein kinase homology  
 C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase  
 F:843-1176/Domain: protein kinase homology <KIN>  
 F:851-859/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 52; DB 2; Length 1363;  
 Best Local Similarity 100.0%; Pred. No. 0.5; Mismatches 0; Indels 0; Gaps 0;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMEIFSL 10  
 |||||  
 DB 1101 GVLMEIFSL 1110

Search completed: November 30, 2004, 07:41:23  
 Job time : 21 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 ; Search time 76.4103 Seconds  
(without alignments)  
75.301 Million cell updates/sec

Title: US-10-008-377a-2  
Perfect score: 52  
Sequence: 1 GVLLMEIFSL 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_02.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	52	100.0	Q9N0K8	Q9N0K8 sus scrofa
2	52	100.0	Q28210	Q28210 bos taurus
3	52	100.0	Q9X554	Q9X554 oryctolagus
4	52	100.0	Q8M123	Q8M123 ovis aries
5	52	100.0	Q9N112	Q9N112 ovis aries
6	52	100.0	Q9N111	Q9N111 ovis aries
7	52	100.0	Q9U8V0	Q9U8V0 capreolus
8	52	100.0	Q9P0V8	Q9P0V8 lampetra re
9	52	100.0	Q8CD05	Q8CD05 mus musculu
10	52	100.0	Q42570	Q42570 xenopus lae
11	52	100.0	Q8CJH1	Q8CJH1 rattus norv
12	52	100.0	Q9I356	Q9I356 coturnix co
13	52	100.0	Q9PUL0	Q9PUL0 brachydanto
14	52	100.0	Q9CWS8	Q9CWS8 mus musculu
15	52	100.0	Q95YV9	Q95YV9 halocynthia
16	52	100.0	Q8TDJ5	Q8TDJ5 homo sapien
17	52	100.0	Q8CJH1	Q8CJH1 rattus norv
18	52	100.0	Q9GSH3	Q9GSH3 halocynthia
19	52	100.0	Q8UVR8	Q8UVR8 mus musculu
20	52	100.0	Q8UVR8	Q8UVR8 mus musculu
21	52	100.0	Q8UVR8	Q8UVR8 mus musculu
22	52	100.0	Q8UVR8	Q8UVR8 mus musculu
23	52	100.0	Q8UVR8	Q8UVR8 mus musculu
24	52	100.0	Q8UVR8	Q8UVR8 mus musculu
25	52	100.0	Q8UVR8	Q8UVR8 mus musculu
26	52	100.0	Q8UVR8	Q8UVR8 mus musculu
27	52	100.0	Q8UVR8	Q8UVR8 mus musculu
28	52	100.0	Q8UVR8	Q8UVR8 mus musculu
29	52	100.0	Q8UVR8	Q8UVR8 mus musculu
30	52	100.0	Q8UVR8	Q8UVR8 mus musculu
31	52	100.0	Q8UVR8	Q8UVR8 mus musculu

32	52	100.0	1356	1	VGR2 HUMAN	P35968 homo sapien
33	52	100.0	1363	1	VGR3 MOUSE	P35917 mus musculu
34	52	100.0	1363	2	Q86M07	Q86M07 homo sapien
35	52	100.0	1367	2	Q912T1	Q912T1 rattus norv
36	52	100.0	1367	1	VGR2 MOUSE	P35918 mus musculu
37	52	100.0	1379	2	F79701	F79701 coturnix co
38	52	100.0	1620	1	ALK HUMAN	Q9UM73 homo sapien
39	52	100.0	1621	2	ALK MOUSE	P97793 mus musculu
40	51	98.1	143	2	Q60515	Q60515 mus musculu
41	51	98.1	207	2	Q9P8G8	Q9P8G8 gallus gall
42	51	98.1	345	2	Q9P0V7	Q9P0V7 lampetra re
43	51	98.1	395	2	Q70W10	Q70W10 cliona intes
44	51	98.1	395	2	CAD58833	CAD58833 cliona int
45	51	98.1	457	2	Q90269	Q90269 brachydanto

## ALIGNMENTS

```

RESULT 1
ID Q9N0K8      PRELIMINARY;      PRT;      116 AA.
AC Q9N0K8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Flc-1 type VEGF receptor (Fragment).
GN Name=Flc-1; (pig).
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oviduct;
RA Kettler A., Mollenhaupt K., Einspanier R.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ245445; CAB76366.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004713; F:Protein-tyrosine kinase activity; IEA.
DR GO; GO:0004872; F:Receptor activity; IEA.
DR GO; GO:0006468; P:Protein amino acid phosphorylation; IEA.
DR InterPro; IP001105; Kinase like.
DR InterPro; IP000719; Prot_Kinase.
DR InterPro; IP001245; Tyr_Kinase.
DR Pfam; PF00069; Pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot_Kinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Receptor.
FT NON_TER      1
FT NON_TER      1
SQ SEQUENCE      116 AA; 13289 MW; 4663DE769F50FA72 CRC64;
Query Match      100.0%; Score 52; DB 2; Length 116;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1 GVLLMEIFSL 10
DB 9 GVLLMEIFSL 18

RESULT 2
ID Q28210      PRELIMINARY;      PRT;      153 AA.
AC Q28210;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE VEGF-Receptor (Flc-Receptor) (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovinae; Bos.  
 OK NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=96212225; PubMed=8626709;  
 RA Mandioca S.J., Menoud P., Pepper M.S.;  
 RT "Transforming growth factor beta 1 downregulates vascular endothelial  
 growth factor receptor-2/flk-1 expression in vascular endothelial  
 cells.";  
 RL J Biol. Chem. 271:11500-11505(1996).  
 DR EMBL X84263; CAA63939.1;  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0004687; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR011009; Kinase like.  
 DR InterPro: IPR000719; Prot kinase.  
 DR Pfam PF00069; PKinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot kinase; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KM Kinase; Receptor; Transferrase; Tyrosine-protein kinase.  
 FT NON\_TER 1 178 178  
 RN [1]

Query Match 100.0%; Score 52; DB 2; Length 153;  
 Best Local Similarity 100.0%; Pred. No. 0.23;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10  
 DB 20 GVLMWEIFSL 29

RESULT 3  
 Q9X554 PRELIMINARY; PRT; 178 AA.  
 ID Q9X554  
 AC Q9X554;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE VEGF receptor-2/flk-1 (Fragment).  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OK NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Japanese white; TISSUE=Liver;  
 RA Uneyki K., Kon K., Ohtaki S.;  
 RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL AB017155; BAA76520.1;  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0004687; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR011009; Kinase like.  
 DR InterPro: IPR000719; Prot kinase.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR InterPro: IPR008266; Tyr\_kinase\_AS.  
 DR Pfam PF00069; PKinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot kinase; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KM Kinase; Receptor; Transferrase; Tyrosine-protein kinase.  
 FT NON\_TER 1 178 178  
 RN [1]

SEQ SEQUENCE 178 AA; 20363 MW; B5F1D07A2AD0A477 CRC64;  
 Query Match 100.0%; Score 52; DB 2; Length 178;  
 Best Local Similarity 100.0%; Pred. No. 0.26;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10  
 DB 74 GVLMWEIFSL 83

RESULT 4  
 Q8M123 PRELIMINARY; PRT; 180 AA.  
 ID Q8M123  
 AC Q8M123;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Vascular endothelial growth factor receptor-2 (Fragment).  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Caprinae; Ovis.  
 OK NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placental artery endothelium;  
 RA Chung J.-Y., Tsai S.C.M., Wen Y.-X., Magness R.R., Zhang J.;  
 RL Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL AF534634; XAN04105.1;  
 DR HSP; P35968; IPR2.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0004687; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR011009; Kinase like.  
 DR InterPro: IPR000719; Prot kinase.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR InterPro: IPR008266; Tyr\_kinase\_AS.  
 DR Pfam PF00069; PKinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot kinase; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KM Kinase; Receptor; Transferrase; Tyrosine-protein kinase.  
 FT NON\_TER 1 180 180  
 FT NON\_TER 180 180  
 SEQ SEQUENCE 180 AA; 20599 MW; 70E5F444574779A0 CRC64;  
 Query Match 100.0%; Score 52; DB 2; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 0.27;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10  
 DB 74 GVLMWEIFSL 83

RESULT 5  
 Q8N112 PRELIMINARY; PRT; 207 AA.  
 ID Q8N112  
 AC Q8N112;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE KDR/flk-1 (Fragment).  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Caprinae; Ovis.  
 OK NCBI\_TaxID=9940;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA Scholz T.D., Segar J.L.;  
 RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF233076; AAF60280.1; -  
 DR GO: GO:0005524; P:ATP binding; IEA.  
 DR GO: GO:0004713; P:protein-tyrosine kinase activity; IEA.  
 DR GO: GO:0016740; P:transferase activity; IEA.  
 DR GO: GO:000521; P:vascular endothelial growth factor receptor; IEA.  
 DR GO: GO:0005468; P:protein amino acid phosphorylation; IEA.  
 DR GO: GO:0005468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR011009; Kinase\_like.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR01245; Tyr\_kinase.  
 DR InterPro: IPR008266; Tyr\_kinase.  
 DR Pfam: PF00069; Kinase; 1.  
 DR PRINTS: PRO0109; TYRKINASE.  
 DR PRODOM: PD000001; Prot\_kinase; 1.  
 DR PROSITE: PS0011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR KINASE: Tyrosine-protein kinase.  
 FT NON TER 1  
 FT 207  
 SQ SEQUENCE 207 AA; 23691 MW; 51C3950D82F988FC CRC64;

Query Match 100.0%; Score 52; DB 2; Length 207;  
 Best Local Similarity 100.0%; Pred. No. 0.31;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVLMWEIFSL 10  
 Db 81 GVLMWEIFSL 90

RESULT 6  
 Q9N111 PRELIMINARY; PRT; 221 AA.  
 ID Q9N111  
 AC Q9N111  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE P1r-1 (Fragment).  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Caprinae; Ovis;  
 OC NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Scholz T.D., Segar J.L.;  
 RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF233077; AAF60281.1; -  
 DR GO: GO:0005524; P:ATP binding; IEA.  
 DR GO: GO:0004713; P:protein-tyrosine kinase activity; IEA.  
 DR GO: GO:0016740; P:transferase activity; IEA.  
 DR GO: GO:0005468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR011009; Kinase\_like.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR01245; Tyr\_kinase.  
 DR InterPro: IPR008266; Tyr\_kinase.  
 DR Pfam: PF00069; Kinase; 1.  
 DR PRINTS: PRO0109; TYRKINASE.  
 DR PRODOM: PD000001; Prot\_kinase; 1.  
 DR PROSITE: PS0011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR KINASE: Tyrosine-protein kinase.  
 FT NON TER 1  
 FT 221  
 SQ SEQUENCE 221 AA; 25356 MW; 5F3059E2C3773B50 CRC64;

Query Match 100.0%; Score 52; DB 2; Length 221;  
 Best Local Similarity 100.0%; Pred. No. 0.33;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVLMWEIFSL 10  
 Db 81 GVLMWEIFSL 90

## RESULT 7

ID Q98V0 PRELIMINARY; PRT; 308 AA.  
 AC Q98V0  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Kit-like A (Fragment).  
 OS Eptatretus burgeti (Inshore hagfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Hyperoartri; Myxiniiformes;  
 OC Myxiniidae; Eptatretinae; Eptatretus.  
 OC NCBI\_TaxID=7764;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20020330; PubMed=10552041;  
 RA Suga H., Hoshiyama D., Kuraku S., Katoh K., Kubokawa K., Miyata T.;  
 RT "Protein tyrosine kinase CDNA from amphioxus, hagfish, and lamprey;  
 RT isoform duplications around the divergence of cyclostomes and  
 RT gnathostomes";  
 RL J. Mol. Evol. 49:601-608(1999).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
 CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. CSF-  
 CC 1/PDGFR receptor subfamily.  
 DR EMBL: AB025553; BAA84743.1; -  
 DR GO: GO:0016020; C:membrane; IEA.  
 DR GO: GO:0005524; P:ATP binding; IEA.  
 DR GO: GO:0004714; P:transmembrane receptor protein tyrosine kin. . .; IEA.  
 DR GO: GO:0005468; P:protein amino acid phosphorylation; IEA.  
 DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; IEA.  
 DR InterPro: IPR011009; Kinase\_like.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR01824; ReceptTykKinIII.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR InterPro: IPR008266; Tyr\_kinase.  
 DR Pfam: PF00069; Kinase; 1.  
 DR PRODOM: PD000001; Prot\_kinase; 1.  
 DR SMART: SMO0219; TyrcK; 1.  
 DR PROSITE: PS0011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 DR KINASE: Tyrosine-protein kinase; Receptor; Transferase; Transmembrane;  
 KM Tyrosine-protein kinase.  
 FT NON TER 1  
 FT 308  
 SQ SEQUENCE 308 AA; 34459 MW; FC63206DC6E629B8 CRC64;

Query Match 100.0%; Score 52; DB 2; Length 308;  
 Best Local Similarity 100.0%; Pred. No. 0.45;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVLMWEIFSL 10  
 Db 233 GVLMWEIFSL 242

## RESULT 8

ID Q9PV08 PRELIMINARY; PRT; 314 AA.  
 AC Q9PV08  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Kit-like A (Fragment).  
 OS Lampetra reissneri (Far Eastern brook lamprey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;  
 OC Petromyzontiformes; Petromyzontidae; Lethenteron.  
 OC NCBI\_TaxID=7753;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=20020330; PubMed=10552041;  
 RA Suga H., Hoshiyama D., Kuraku S., Katch K., Kubokawa K., Miyata T.,  
 RT "Protein tyrosine kinase CDNs from amphioxus, hagfish, and lamprey;  
 RT isoform duplications around the divergence of cyclostomes and  
 RT gnathostomes.";  
 RL J. Mol. Evol. 49:601-608(1999).  
 DR EMBL; AB025555; BAA84745.1; -  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR011009; Kinase\_like.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR InterPro; IPR008266; Tyr\_kinase\_AS.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR ProDom; PD000001; Prot\_kinase; 2.  
 DR SMART; SM00219; TYR\_K1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KM Kinase; Transferase; Tyrosine-protein kinase.  
 FT NON\_TER  
 SQ SEQUENCE 314 AA; 35254 MW; D375DB78DDADFC CRC64;  
 Query Match 100.0%; Score 52; DB 2; Length 314;  
 Best Local Similarity 100.0%; Pred. No. 0.46;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GVLMEIFSL 10  
 DB 243 GVLMEIFSL 252

RESULT 9  
 ID Q8CD05 PRELIMINARY; PRT; 341 AA.  
 AC Q8CD05  
 DT 01-MAR-2003 (TREMblrel. 23, Created)  
 DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)  
 DE Mus musculus 14 days embryo thymus cDNA, RIKEN full-length enriched  
 DE library, clone6130401C07 product:kinase insert domain protein  
 DE receptor, full insert sequence. (Fragment).  
 GN Name:Kdr;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RX MEDLINE=9279253; PubMed=10349636;  
 RA Carninci P., Hayashizaki Y.,  
 RT "High-efficiency full-length cDNA cloning.";  
 RT Meth. Enzymol. 303:19-44(1999).  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RA the FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RX MEDLINE=20499374; PubMed=11042159;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.,  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RX MEDLINE=20530913; PubMed=11076861;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Kono H., Akiyama J., Nishi K., Katsunari T., Tashiro H., Itoh M.,  
 RA Suni N., Ishi Y., Nakamura S., Hazama M., Nishize T., Harada A.,  
 RA Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kasaiwaig K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Matsubiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
 RA Hori F., Imotani K., Ishi Y., Itoh M., Kagawa I., Kasukawa T.,  
 RA Kachi H., Kawai J., Kojima Y., Kondo S., Kono H., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Muzra M., Nakamura M.,  
 RA Nishi K., Nomura K., Numazaki R., Ono M., Onato N., Okazaki Y.,  
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sato H.,  
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK031739; BAC27532.1; -  
 DR HSEF; P35968; IWR2.  
 DR MCD; MGI:36683; Kdr.  
 DR GO; GO:0045165; P:cell fate commitment; IMP.  
 DR GO; GO:0045446; P:endothelial cell differentiation; IDA.  
 DR InterPro; IPR001009; Kinase\_like.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR InterPro; IPR008266; Tyr\_kinase\_AS.  
 DR ~InterPro; IPR009136; VEGFR2.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR PRINTS; PR01834; VEGFRECEPTR2.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KM Kinase; Receptor; Transferase; Tyrosine-protein kinase.  
 FT NON\_TER  
 SQ SEQUENCE 341 AA; 38302 MW; E2B4DCC4BB481195 CRC64;  
 Query Match 100.0%; Score 52; DB 2; Length 341;  
 Best Local Similarity 100.0%; Pred. No. 0.5;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GVLMEIFSL 10  
 DB 86 GVLMEIFSL 95

RESULT 10  
 ID Q42570 PRELIMINARY; PRT; 346 AA.  
 AC Q42570  
 DT 01-JAN-1998 (TREMblrel. 05, Created)  
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)  
 DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)  
 DE Receptor tyrosine kinase (Fragment).

```

GN Name=flx-1;
OS Xenopus laevis (African clawed frog);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodidae; Xenopus.
NCBI_Taxid=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Head;
RT "Neovascularization of the Xenopus embryo.";
RL Dev. Dyn. 0:0-0(1997).
DR EMBL; AF007760; F:ATP binding; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0005021; F:vascular endothelial growth factor receptor; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR InterPro; IPR009136; VEGFR2.
DR Pfam; PF00069; PKinase; 1.
DR PRINTS; PR0109; TYRKINASE.
DR PRINTS; PR01834; VEGFRECEPTR2.
DR PRODOM; PD000001; Prot_kinase; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KM Kinase; Receptor; Transferase; Tyrosine-protein kinase.
FT NON_TER 1
FT NON_TER 346
SQ SEQUENCE 346 AA; 38976 MW; AA35FA8F0C702416 CRC64;

Query Match 100.0%; Score 52; DB 2; Length 346;
Best Local Similarity 100.0%; Pred. No. 0.51;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10
DB 88 GVLMWEIFSL 97

RESULT 11
Q8CJH1 PRELIMINARY; PRT; 351 AA.
AC Q8CJH1;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2004 (TEMBLrel. 25, Last sequence update)
DE Anaplastic lymphoma kinase (Fragment).
GN Name=alk;
OS Rattus norvegicus (Rat);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RA Asahina K., Sato H., Yamasaki C., Tateno C., Kataoka M., Shikawa M.,
RA Katayama S., Yoshizato K.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB073169; BAC21663.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002011; RecepttyrkineII.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.

```

```

DR Pfam; PF00069; PKinase; 1.
DR PRINTS; PR0109; TYRKINASE.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00219; TYRKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; UNKNOWN_1.
KM ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.
FT NON_TER 1
FT NON_TER 351
SQ SEQUENCE 351 AA; 38984 MW; 6ADCEC220C76B125 CRC64;

Query Match 100.0%; Score 52; DB 2; Length 351;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10
DB 319 GVLMWEIFSL 328

RESULT 12
Q91356 PRELIMINARY; PRT; 367 AA.
AC Q91356;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-OCT-1996 (TEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Endothelial kinase Quek2 (Fragment).
GN Name=endothelial kinase Quek2;
OS Corvus corax (Common quail);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Cuculix.
NCBI_Taxid=9091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93378866; PubMed=8396413;
RT Eichmann A., Marcelle C., Breant C., le Douarin N.M.;
RT "Two molecules related to the VEGF receptor are expressed in early
RT endothelial cells during avian embryonic development.";
RL Mech. Dev. 42:33-48(1993).
DR EMBL; S65207; AAB28128.1; -.
DR PIR; B56598; B56598.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF00069; PKinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot_kinase; 1.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KM Kinase; Transferase; Tyrosine-protein kinase.
FT NON_TER 367
FT NON_TER 367
SQ SEQUENCE 367 AA; 41792 MW; 04442B35940077A0 CRC64;

Query Match 100.0%; Score 52; DB 2; Length 367;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10
DB 101 GVLMWEIFSL 110

RESULT 13
Q9PULO

```

ID Q9PUL0 PRELIMINARY; PRT; 404 AA.  
 AC Q9PUL0  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE F14 protein (Fragment).  
 GN Name-F14;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OC NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98294174; PubMed=9630750;  
 RA Thompson M.A., Ransom D.G., Pratt S.J., MacLennan H., Kieran M.W.,  
 RA Derrich H.W., I.L.I., Vail B., Huber T.L., Paw B., Brownlie A.J.,  
 RA Oates A.C., Fritz A., Gates M.A., Amos A., Bahary N., Talbot W.S.,  
 RA Her H., Beler D.R., Postlethwait J.H., Zon L.I.;  
 RA EMBL; AF178759; AAD56011.2; -  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 DR ZFIN; ZDB-GENE-980526-326; f14.  
 DR ZFIN; ZDB-GENE-980526-326; f14.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR011003; Kinase\_like.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR InterPro; IPR008266; Tyr\_kinase\_AS.  
 DR Pfam; PF00069; Kinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR PROSITE; PSS0011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PSS0109; PROTEIN\_KINASE\_TYR; 1.  
 DR KINASE; transferase; Tyrosine-Protein kinase.  
 FT NON\_TER 1  
 SQ SEQUENCE 404 AA; 45281 MW; A149B836219EB4A9 CRC64;  
 Query March 100.0%; Score 52; DB 2; Length 404;  
 Best Local Similarity 100.0%; Pred. No. 0.59; 0; Indels 0; Gaps 0;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GVLLWETFSL 10  
 Db 156 GVLLWETFSL 165  
 RESULT 14  
 Q9CWS8 PRELIMINARY; PRT; 453 AA.  
 AC Q9CWS8  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Mus musculus adult male lung cDNA, RIKEN full-length enriched library,  
 DE clone:1200012B22 product:FMS-like tyrosine kinase 1, full insert  
 DE sequence. (Fragment).  
 GN Name-F14;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RX MEDLINE=99279253; PubMed=10349636;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA RIKEN PANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RA The PANTOM Consortium  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RX MEDLINE=20499374; PubMed=11042159;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RX MEDLINE=20510913; PubMed=11076661;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Kono H., Akiyama J., Nishi K., Kusunagi T., Teshiro H., Itoh M.,  
 RA Sato N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,  
 RA Yamamoto R., Matsunoto H., Sakauechi S., Ikegami T., Kasaiwagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Matsubiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
 RA Arikawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hirooka T., Hori F.,  
 RA Imclani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
 RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,  
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ono M.,  
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai T.,  
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
 RA Sogabe Y., Suzuki H., Tagami M., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
 RA Muramatsu M., Hayashizaki Y.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK004721; BAB23504.2; -  
 DR HSRP; P35868; ITR2  
 DR MGD; MGI:95858; F14.1  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR011003; Kinase\_like.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR InterPro; IPR008266; Tyr\_kinase\_AS.  
 DR Pfam; PF00069; Kinase; 2.  
 DR ProDom; PD000001; Prot\_kinase; 2.  
 DR SMART; SMO0219; TyKc; 1.

DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00109; PROTEIN KINASE\_TYR; 1.  
 KM Kinase; Transferase; Tyrosine-protein kinase.  
 FT NON TER 1  
 SQ SEQUENCE 453 AA; 51239 MW; 2A48AD0DC677535 CRC64;

Query Match 100.0%; Score 52; DB 2; Length 453;  
 Best Local Similarity 100.0%; Pred. No. 0.66;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
 |||||  
 Db 206 GVLWEIFSL 215

RESULT 15  
 ID Q95YM9 PRELIMINARY; PRT; 763 AA.  
 AC Q95YM9;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE RCFR.  
 GN Name=Hrfgfr.  
 OS Halocynthia roretzi (Sea squirt).  
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;  
 OC Stolidobranchia; Pyridae; Halocynthia.  
 OX NCBI\_TaxID=7729;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2147207; PubMed=11526077;  
 RA Shimauchi Y., Murakami S.D., Satoh N.;  
 RT "Fgf signals are involved in the differentiation of notochord cells  
 and mesenchyme cells of the ascidian Halocynthia roretzi.";  
 RL Development 128:2711-2721(2001).  
 DR EMBL; AB046873; BAB59007.1; -.  
 DR HSSP; P11362; 1EVT.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR InterPro; IPR011009; Kinase\_like.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR InterPro; IPR008266; Tyr\_kinase\_AS.  
 DR Pfam; PF00047; Ig\_2.  
 DR Pfam; PF00069; Kinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR SMART; SM00408; IGC2; 2.  
 DR SMART; SM00219; TYRC; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 2.  
 DR PROSITE; PS50011; PROTEIN KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN KINASE\_TYR; 1.  
 KM Kinase; Transferase; Tyrosine-protein kinase.  
 SQ SEQUENCE 763 AA; 86741 MW; A34C1871DB796950 CRC64;

Query Match 100.0%; Score 52; DB 2; Length 763;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
 |||||  
 Db 601 GVLWEIFSL 610

Search completed: November 30, 2004, 07:17:27  
 Job time : 77.4103 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 ; Search time 77.1795 Seconds  
(Without alignments)  
46.460 Million cell updates/sec

Title: US-10-008-377a-2  
Perfect score: 52  
Sequence: 1 GVLTMEIFSL 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: Genesep235ep04:\*  
2: Genesep1980s:\*  
3: Genesep1990s:\*  
4: Genesep2000s:\*  
5: Genesep2001s:\*  
6: Genesep2002s:\*  
7: Genesep2003as:\*  
8: Genesep2003bs:\*  
9: Genesep2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	10	-6	AAO22984
2	52	100.0	116	4	AAO22984 Human p37
3	52	100.0	116	4	AAO22984 Human ALK
4	52	100.0	116	4	AAO22984 Leukocyte
5	52	100.0	116	6	ABG74477 Tyrosine
6	52	100.0	116	6	ABG74481 Tyrosine
7	52	100.0	127	4	ABBI1041 Human pro
8	52	100.0	316	3	AAO22984 Human VEG
9	52	100.0	348	2	AAO22984 Protein t
10	52	100.0	366	3	AAO22984 Human VEG
11	52	100.0	525	2	AAO22984 NPW/ALK f
12	52	100.0	525	6	AAO22984 Human exp
13	52	100.0	572	2	AAO22984 Partial A
14	52	100.0	675	4	ABG20950 Novel hum
15	52	100.0	680	4	AAO22984 Human NPM
16	52	100.0	680	5	ABG95067 Human tra
17	52	100.0	680	6	ABU04362 Human exp
18	52	100.0	680	6	ABU04362 Human exp
19	52	100.0	680	6	ABU04362 Human exp
20	52	100.0	713	7	ABG74472 NPW/ALK f
21	52	100.0	713	7	ABG74472 NPW/ALK f
22	52	100.0	771	8	ADOC4223 Human VEG
23	52	100.0	832	8	ADH48374 Human kin
24	52	100.0	1273	3	AAO22984 Human KRP
25	52	100.0	1298	2	AAO22984 Human tyr

26	52	100.0	1298	2	AAO22984	AAO22984
27	52	100.0	1298	2	AAO22984	AAO22984
28	52	100.0	1298	3	AAO22984	AAO22984
29	52	100.0	1298	3	AAO22984	AAO22984
30	52	100.0	1298	3	AAO22984	AAO22984
31	52	100.0	1298	3	AAO22984	AAO22984
32	52	100.0	1298	3	AAO22984	AAO22984
33	52	100.0	1298	5	ABG32043	ABG32043
34	52	100.0	1298	6	ABU07853	ABU07853
35	52	100.0	1298	6	ABU04292	ABU04292
36	52	100.0	1298	6	ABU04307	ABU04307
37	52	100.0	1298	6	ABU04298	ABU04298
38	52	100.0	1298	6	ABU04304	ABU04304
39	52	100.0	1298	6	ABU04296	ABU04296
40	52	100.0	1298	6	ABU04277	ABU04277
41	52	100.0	1298	6	ABU04274	ABU04274
42	52	100.0	1298	6	ABU04283	ABU04283
43	52	100.0	1298	6	ABU04284	ABU04284
44	52	100.0	1298	6	ABU04293	ABU04293
45	52	100.0	1298	7	ADD08956	ADD08956

## ALIGNMENTS

RESULT 1  
AAO22984  
ID AAO22984 standard; peptide; 10 AA.  
AC AAO22984;  
XX  
XX 17-SEP-2003 (first entry)  
XX  
XX Human p376-85 ALK-derived HLA-A\*0201 restricted CTL epitope peptide.  
DE HLA-A\*0201 binding; anaplastic lymphoma kinase; ALK; CTL epitope; ALCL;  
XX cytotoxic T-lymphocyte; receptor tyrosine; NPW/ALK fusion protein; human;  
XX oncogenic; anaplastic large cell lymphoma; neuroblastoma; gene therapy;  
XX t(2;5) translocation; neoplasia; cytostatic; immunotherapy; alkoma;  
XX p376-85.  
XX Homo sapiens.  
OS  
XX WO2003042243-A2.  
XX  
XX 22-MAY-2003.  
XX  
XX 14-NOV-2002; 2002NO-EP012764.  
XX  
XX 15-NOV-2001; 2001US-0008377.  
XX  
XX (NAST-) 1ST NAZ STUDIO E CURA DEI TUMORI.  
XX (SUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
XX  
XX Gambacorti-Passerini C, Passoni L;  
XX WPI, 2003-441791/41.  
XX  
XX New HLA-A\*0201-binding Anaplastic Lymphoma Kinase (ALK) peptide,  
XX useful for preparing a composition for treating ALK-positive lymphoma,  
XX neuroblastoma or ALK-expressing neoplasia.  
XX  
XX Claim 1, Page 7, 33pp; English.  
XX  
XX The invention relates to a novel HLA-A\*0201-binding anaplastic lymphoma  
XX kinase (ALK)-derived peptide, of 9-10 amino acids, which may act as a  
XX cytotoxic T-lymphocyte (CTL)-epitope. ALK is a receptor tyrosine kinase  
XX which forms part of the NPW/ALK fusion protein (oncogenic lymphoma kinase  
XX fusion protein) in ALCL (anaplastic large cell lymphoma)-derived cell  
XX lines. More than 50% of ALCL cases possess a t(2;5) chromosomal  
XX translocation that leads to the expression of the NPW/ALK fusion protein  
XX which forms a potent oncogene when constitutively activated. Translocated  
XX ALK is a widely expressed tumour-associated antigen characteristic of ALK

CC -positive lymphomas, neuroblastomas and ALK-expressing neoplasias. The  
 CC peptides of the invention demonstrate cytostatic activity and induce an  
 CC MHC (major histocompatibility complex) class I restricted cytotoxic  
 CC lymphocyte response against tumor cells expressing the NPM/ALK fusion  
 CC protein. Hence, the peptides may be utilized during the treatment, via  
 CC immunotherapy, of ALK-positive lymphomas such as alkomas, neuroblastomas  
 CC and ALK-expressing neoplasias. In addition, the peptides may be used  
 CC during gene therapy. The current sequence is that of the human p376-85  
 CC ALK-derived HLA-A\*0201 restricted CTL epitope peptide of the invention  
 XX

Sequence 10 AA:

Query Match 100.0%; Score 52; DB 6; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.017;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLLWEIFSL 10  
 DB 1 GVLLWEIFSL 10

RESULT 2  
 AAB73863  
 ID AAB73863 standard; protein; 116 AA.

AC AAB73863;

DT 15-MAY-2001 (first entry)

DE Human ALK catalytic domain sequence #4.

KW Human; anaplastic lymphoma kinase; ALK; nucleophosmin; NPM;

KM Chromosomal translocation; cancer; NPM/ALK; catalytic domain.

XX Homo sapiens.

OS US6174674-B1.

PN 16-JAN-2001.

PF 19-JUN-1998; 98US-00100089.

PR 03-DEC-1993; 93US-00160861.

PT 12-OCT-1995; 95US-00542363.

PS (SQUID-) ST JUDE CHILDREN'S RES HOSPITAL.

PI Morris SW, Look AT;

DR WPI; 2001-243208/25.

XX Detection of chromosomal rearrangement or translocations present in t(2;

PT 5) for cancer diagnosis comprises using and identifying anaplastic

PT lymphoma kinase and nucleophosmin genes in nucleic acid hybridization

PT techniques.

PS Example 2; Fig 2C; 87pp; English.

XX The present sequence is given in a specification relating to a method for

CC detecting a chromosomal rearrangement involving a breakpoint in the

CC anaplastic lymphoma kinase (ALK) or nucleophosmin (NPM) gene. The method

CC is based on the identification and sequence determination of a novel

CC gene, ALK, which is fused to the gene encoding NPM in translocations

CC present in t(2/5) lymphoma cells. The method is useful for detecting a

CC t(2/5) chromosomal rearrangement, or a chromosomal rearrangement

CC involving the ALK gene or NPM gene. It is particularly useful in

CC determining if particular cells or tissues express ALK or NPM/ALK coding

CC sequences. The method is also useful in diagnostic assays to determine,

CC for example, if a mammal has cancer or a genetic predisposition to cancer

XX Sequence 116 AA;

SO Query Match 100.0%; Score 52; DB 4; Length 116;

Best Local Similarity 100.0%; Pred. No. 0.22;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLLWEIFSL 10  
 DB 56 GVLLWEIFSL 65

RESULT 3  
 AAB73867  
 ID AAB73867 standard; protein; 116 AA.

AC AAB73867;

DT 15-MAY-2001 (first entry)

DE Leukocyte tyrosine kinase catalytic domain sequence #4.

KW Leukocyte tyrosine kinase; LTK; protein tyrosine kinase; PTK;

KM insulin receptor subfamily; anaplastic lymphoma kinase; ALK;

KM nucleophosmin; NPM; chromosomal translocation; cancer; NPM/ALK;

XX catalytic domain.

OS Unidentified.

PN US6174674-B1.

PD 16-JAN-2001.

PF 19-JUN-1998; 98US-00100089.

PR 03-DEC-1993; 93US-00160861.

PT 12-OCT-1995; 95US-00542363.

PS (SQUID-) ST JUDE CHILDREN'S RES HOSPITAL.

PI Morris SW, Look AT;

DR WPI; 2001-243208/25.

XX Detection of chromosomal rearrangement or translocations present in t(2;

PT 5) for cancer diagnosis comprises using and identifying anaplastic

PT lymphoma kinase and nucleophosmin genes in nucleic acid hybridization

PT techniques.

PS Example 2; Fig 2C; 87pp; English.

XX The present sequence is given in a specification relating to a method for

CC detecting a chromosomal rearrangement involving a breakpoint in the

CC anaplastic lymphoma kinase (ALK) or nucleophosmin (NPM) gene. The method

CC is based on the identification and sequence determination of a novel

CC gene, ALK, which is fused to the gene encoding NPM in translocations

CC present in t(2/5) lymphoma cells. The method is useful for detecting a

CC t(2/5) chromosomal rearrangement, or a chromosomal rearrangement

CC involving the ALK gene or NPM gene. It is particularly useful in

CC determining if particular cells or tissues express ALK or NPM/ALK coding

CC sequences. The method is also useful in diagnostic assays to determine,

CC for example, if a mammal has cancer or a genetic predisposition to cancer

XX Sequence 116 AA;

SO Query Match 100.0%; Score 52; DB 4; Length 116;

Best Local Similarity 100.0%; Pred. No. 0.22;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLLWEIFSL 10

DB 56 GVLLWEIFSL 65

RESULT 4

ABG74477

ID ABG74477 standard; protein; 116 AA.

XX ABG74477;  
AC  
XX  
XX  
DT 11-APR-2003 (first entry)  
XX  
XX Tyrosine kinase insulin receptor ALK fragment SEQ ID 15.  
DE  
XX  
XX ALK; human; nucleophosmin; NPM; anaplastic lymphoma kinase; chromosome 5;  
KM t(2; 5) lymphoma cell; chromosomal rearrangement; chromosome 2;  
KM nucleolar phosphoprotein; centromeric; telomeric.  
XX  
XX Homo sapiens.  
OS  
XX  
XX US6451997-B1.  
PN  
XX  
XX 17-SEP-2002.  
PD  
XX  
XX 28-SEP-2000; 2000US-00670827.  
PF  
XX  
XX 03-DEC-1993; 93US-00160861.  
PR 12-OCT-1995; 95US-00542363.  
PR 19-JUN-1998; 98US-00100089.  
XX  
XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
PA  
XX  
XX Morris SW, Look AT;  
PI  
XX  
XX WPI; 2003-101739/09.  
DR  
XX  
XX Kit for use in method of detecting t(2;5) chromosomal rearrangements or  
PT rearrangements involving nucleolar phosphoprotein (NPM) or anaplastic  
PT lymphoma kinase (ALK) comprises probes that hybridizes with NPM and ALK  
PT genes.  
XX  
XX Example 1; Fig 2C; 87pp; English.  
PS  
XX  
XX This invention describes a novel kit for use in a method of detecting t(2  
CC ;5) chromosomal rearrangements (CR) or CRs involving nucleolar  
CC phosphoprotein (NPM) and anaplastic lymphoma kinase (ALK). The method  
CC comprises labeled probes that hybridize to a sequence of 2442 or 6226  
CC nucleotides or their complement, and to region of human chromosome 5/2  
CC having wild-type NPM/ALK genes (the region is centromeric/telomeric to  
CC nucleotide). This sequence represents a polypeptide sequence described in  
CC the disclosure of the invention  
XX  
XX Sequence 116 AA;  
SQ  
XX  
XX Query Match 100.0%; Score 52; DB 6; Length 116;  
XX Best Local Similarity 100.0%; Pred. No. 0.22;  
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GVLTWEIFSL 10  
DB 56 GVLTWEIFSL 65  
XX  
XX RESULT 5  
XX ABG74481 standard; protein; 116 AA.  
ID  
XX  
XX ABG74481;  
AC  
XX  
XX 11-APR-2003 (first entry)  
DT  
XX  
XX Tyrosine kinase insulin receptor LTK fragment SEQ ID 19.  
DE  
XX  
XX ALK; human; nucleophosmin; NPM; anaplastic lymphoma kinase; chromosome 5;  
KM t(2; 5) lymphoma cell; chromosomal rearrangement; chromosome 2;  
KM nucleolar phosphoprotein; centromeric; telomeric.  
XX  
XX Homo sapiens.  
OS  
XX  
XX US6451997-B1.  
PN

XX 17-SEP-2002.  
PD  
XX  
XX 28-SEP-2000; 2000US-00670827.  
PF  
XX  
XX 03-DEC-1993; 93US-00160861.  
PR 12-OCT-1995; 95US-00542363.  
PR 19-JUN-1998; 98US-00100089.  
XX  
XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
PA  
XX  
XX Morris SW, Look AT;  
PI  
XX  
XX WPI; 2003-101739/09.  
DR  
XX  
XX Kit for use in method of detecting t(2;5) chromosomal rearrangements or  
PT rearrangements involving nucleolar phosphoprotein (NPM) or anaplastic  
PT lymphoma kinase (ALK) comprises probes that hybridizes with NPM and ALK  
PT genes.  
XX  
XX Example 1; Fig 2C; 87pp; English.  
PS  
XX  
XX This invention describes a novel kit for use in a method of detecting t(2  
CC ;5) chromosomal rearrangements (CR) or CRs involving nucleolar  
CC phosphoprotein (NPM) and anaplastic lymphoma kinase (ALK). The method  
CC comprises labeled probes that hybridize to a sequence of 2442 or 6226  
CC nucleotides or their complement, and to region of human chromosome 5/2  
CC having wild-type NPM/ALK genes (the region is centromeric/telomeric to  
CC nucleotide). This sequence represents a polypeptide sequence described in  
CC the disclosure of the invention  
XX  
XX Sequence 116 AA;  
SQ  
XX  
XX Query Match 100.0%; Score 52; DB 6; Length 116;  
XX Best Local Similarity 100.0%; Pred. No. 0.22;  
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GVLTWEIFSL 10  
DB 56 GVLTWEIFSL 65  
XX  
XX RESULT 6  
XX ABB11041 standard; peptide; 127 AA.  
ID  
XX  
XX ABB11041;  
AC  
XX  
XX 11-JAN-2002 (first entry)  
DT  
XX  
XX Human protein-Tyr kinase homologue, SEQ ID NO:1411.  
DE  
XX  
XX Human; cytokine; cell proliferation; cell differentiation; growth factor;  
KM haematopoiesis regulation; tissue growth; immunomodulator; activin;  
KM inhibin; chemokinesis; thrombolysis; oncogenesis;  
KM proliferation; metastasis; cancer; tumor; haematopoietic disorder;  
KM myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
KM chronic inflammatory condition; proliferative retinopathy;  
KM atherosclerosis; coronary heart disease; arterial ischaemia;  
KM bone disorder; osteoporosis; vascular growth disorder;  
KM tissue regeneration; wound healing; infection; immune disorder;  
KM cell culture; drug screening; gene therapy; antiinflammatory;  
KM antiasthmatic; antiarthritis; haemostatic; antiatherosclerotic;  
KM cytosstatic; osteopathic; vasotrophic; cardiac; virucide; antibacterial;  
KM antifungal; vulnery; anticancer.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200157188-A2.  
PN  
XX  
XX 09-AUG-2001.  
PD  
XX  
XX 05-FEB-2001; 2001WO-US003800.  
PF



Best Local Similarity 100.0%; Pred. No. 0.61;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
DB 237 GVLWEIFSL 246

RESULT 8  
ID AAR85938 standard; protein; 348 AA.

XX AAR85938;

DT 14-FEB-1996 (first entry)

DE Protein tyrosine-kinase SAL-SI.

KM Protein tyrosine-kinase; SAL-SI; cell growth; differentiation.

XX Homo sapiens.

XX W09527061-A1.

PD 12-OCT-1995.

PF 04-APR-1995; 95WO-US004228.

PR 04-APR-1994; 94US-00222616.

XX (GETH) GENENTECH INC.

PI Bennett BD, Goeddel D, Lee JM, Matthews W, Tsai SP, Wood WI;

XX WPI; 1995-36160/47.

DR N-PSDB; AAT03101.

XX Agonist antibodies which activate specific protein tyrosine kinase(s) -

PT also activate chimeric proteins of kinase extracellular domain and Ig

XX growth and differentiation.

XX Disclosure; Page 47-48; 125pp; English.

XX DNA probes based on protein tyrosine-kinase (PTK) sequences were used to

CC screen cDNA libraries to identify novel PTK genes. The SAL-SI gene was

CC isolated from several megakaryocytic cell line libraries. The gene (see

CC AAT03101) encodes a protein (AAR85938) showing homology to FLT/FLK family

CC PTKs. SAL-SI can be used to target proteins in drug design, as an

CC (ant)agonist of other PTKs, to screen PTK ligands or to modulate

CC megakaryocyte/platelet adhesion

XX Sequence 348 AA;

QY 1 GVLWEIFSL 10

DB 151 GVLWEIFSL 160

XX AAY87421;

DT 03-JUL-2000 (first entry)

DE Human VEGFR-2 catalytic kinase domain.

KM VEGFR-2; vascular endothelial growth factor receptor-2; angiogenesis;  
KW cytoplasmic domain; kinase insert domain; KID; catalytic region;  
KW receptor tyrosine kinase; X-ray crystallography; drug design.

XX Homo sapiens.  
XX Key Location/Qualifiers  
XX Domain 128..195  
XX /note="Kinase insert domain (KID)"

XX W0200014105-A1.

XX 16-MAR-2000.

PF 07-SEP-1999; 99WO-US020319.

PR 08-SEP-1998; 98US-0099503P.

XX (AGOU-) AGOURON PHARM INC.

PI McFigue MA, Pinko C, Parast CV, Gehring MR, Kan C, Appel K;

PI Wickersham JA, Showalter RE, Tempczyk-Russell A, Mroczkowski B;

PI Villalanza JE;

DR WPI; 2000-256936/22.

XX Assaying a candidate compound for its ability to interact with a modified

PT receptor tyrosine kinase polypeptide, especially vascular endothelial

PT growth factor-2, by using molecular modelling techniques such as X-ray

PT crystallography.

XX Example 1; Fig 1; 47pp; English.

XX The invention relates to methods of assaying candidate drugs for their

CC ability to interact with modified receptor tyrosine kinases (RTKs) via

CC cloning, sequencing and X-ray crystallography. In particular, it relates

CC to the identification of vascular endothelial growth factor receptor-2

CC (VEGFR-2) agonists and antagonists via the use of X-ray crystallographic

CC data obtained from a mutant human VEGFR-2 catalytic cytoplasmic domain,

CC VEGFR-2 50P (AAY87421). The cytoplasmic domain of VEGFR-2, a key enzyme

CC in angiogenesis, has receptor tyrosine kinase (RTK) activity and mediates

CC signal transduction. VEGFR-2 50P comprises residues 606-939 and 990-1171

CC of wild-type human VEGFR-2. Relative to the wild-type VEGFR-2, VEGFR-2

CC 50P sequence has a deletion of 50 amino acids in the kinase insert domain

CC (KID), retaining only 18 amino acids of the KID. The KID is

CC characteristic of the platelet-derived growth factor receptor (PDGFR)

CC family, to which VEGFR-2 belongs. It is not thought to be necessary for

CC intrinsic tyrosine kinase activity, but is important for the binding of

CC other proteins involved in signal transduction via autophosphorylation of

CC KID tyrosine residues. VEGFR-2 50P retains comparable in vitro RTK

CC activity to that of wild-type VEGFR-2 kinase domain (AAY87421), but is

CC crystallizable, allowing the structure of the phosphorylated unliganded

CC VEGFR-2 50P to be determined via X-ray crystallography to a resolution of

CC 2.4 Angstroms. The structure of the VEGFR-2 50P can provide insights into

CC the orientation of the KID of VEGFR-2, which may be relevant to other

CC PDGFR family members. The structural information can be used in molecular

CC modelling to model the interaction of the VEGFR-2 catalytic domain with

CC candidate agonists and antagonists, and can be used in de novo drug

CC design. Compounds identified via the methods of the invention can be used

CC for treating or preventing diabetic retinopathy or Kaposi's sarcoma,

CC suppressing the development of kidney diseases, preventing the joint

CC destruction that accompanies rheumatoid arthritis, suppressing

CC manifestations of psoriasis, suppressing accelerated atherosclerosis,

CC suppressing coronary artery disease and other vasculopathies due to

CC atherogenesis, and suppressing tumour growth via paracrine or autocrine

CC mediated responses to other cytokines. The present sequence represents

CC the wild-type human VEGFR-2 catalytic domain

XX Sequence 366 AA;

QY 366 AA;

DB 366 AA;

XX AAY87421;

DT 03-JUL-2000 (first entry)

DE Human VEGFR-2 catalytic kinase domain.

XX Query Match 100.0%; Score 52; DB 3; Length 366;  
Best Local Similarity 100.0%; Pred. No. 0.71;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVLMWEIFSL 10  
 Db 287 GVLMWEIFSL 296

## RESULT 10

ID AAR78521 standard; protein; 525 AA.

XX AAR78521;

XX 25-MAR-2003 (revised)

DT 15-NOV-1995 (first entry)

XX NPM/ALK fusion protein.

XX ALK protein, protein-tyrosine-kinase; fusion protein; NPM protein;

KW nuclear phosphoprotein; anaplastic large cell lymphoma; t(2;

KW 5) translocation; diagnosis; gene therapy.

XX Homo sapiens.

OS W09515331-A1.

XX 08-JUN-1995.

XX 05-DEC-1994; 94MO-US013947.

XX 03-DEC-1993; 93US-00160861.

XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.

XX Morris SW, Look AT;

XX WPI; 1995-215226/28.

XX N-PSDB; AAQ95542.

XX Methods for detecting human t(2;5) lymphoma - for detection and diagnosis

XX of anaplastic large cell lymphoma(s).

XX Claim 16; Page 42-44; 70pp; English.

XX The translocation event that occurs in human t(2;5) lymphoma brings

CC sequences from the nuclear phosphoprotein (NPM) gene on chromosome 5q35

CC to those from a protein-tyrosine-kinase (ALK) gene on chromosome-2q23.

CC Nucleic acids encoding a partial ALK protein (AAR78520) and the ALK/NPM

CC fusion protein (AAR78521) were isolated (AAQ95541, AAQ95542).

CC Identification of the NPM/ALK fusion allows lymphoma diagnosis, or

CC therapy through the use of antisense RNA or ribozymes. (Updated on 25-MAR

CC -2003 to correct PV field.)

XX Sequence 525 AA;

XX Query Match 100.0%; Score 52; DB 2; Length 525;

XX Best Local Similarity 100.0%; Pred. No. 1;

XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 GVLMWEIFSL 10

XX 376 GVLMWEIFSL 385

KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;

XX protease; protease inhibitor; transporter; cytoskeletal protein;

XX receptor; transcription factor; cancer; MHC;

KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;

XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX Homo sapiens.

OS W0200278524-A2.

XX 10-OCT-2002.

XX 28-MAR-2002; 2002MO-US009671.

XX 28-MAR-2001; 2001US-0279495P.

XX 21-MAY-2001; 2001US-0292544P.

XX 08-AUG-2001; 2001US-0310801P.

XX 01-OCT-2001; 2001US-0326370P.

XX 04-DEC-2001; 2001US-0356780P.

XX 20-FEB-2002; 2002US-0358985P.

XX (ZYCO-) ZYCOG INC.

XX Chicx RM, Tomlinson AJ, Urban RG;

XX WPI; 2003-040607/03.

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,

XX cytoskeletal proteins, receptors or transcription factors), useful for

XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or

XX leukemia.

XX Example 2; SEQ ID NO 1034; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a

CC fragment of a kinase, phosphatase, protease, protease inhibitor,

CC transporter, cytoskeletal protein, receptor or transcription factor. The

CC polypeptide is useful as an immunogenic composition for eliciting in a

CC mammal an immunogenic response directed against any of the purified

CC polypeptide. The purified polypeptide, or the antibody that binds to this

CC polypeptide, is useful for treating cancer. The polypeptide is also

CC useful for identifying compounds that binds to a naturally processed

CC class I or class II MHC-binding polypeptide. The polypeptides and

CC polynucleotides are particularly useful for treating or preventing

CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,

CC lymphoma or leukaemia. These are also useful for screening agents for

CC treating the above mentioned diseases. This sequence represents an

CC expressed protein tag (EPT) isolated from human tissue for translational

CC profiling. Note: This sequence does not appear in the printed

CC specification but was obtained in electronic format directly from WIPO at

XX ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 525 AA;

XX Query Match 100.0%; Score 52; DB 6; Length 525;

XX Best Local Similarity 100.0%; Pred. No. 1;

XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 GVLMWEIFSL 10

XX 376 GVLMWEIFSL 385

## RESULT 12

ID AAR78520 standard; protein; 572 AA.

XX AAR78520;

XX 25-MAR-2003 (revised)

DT 15-NOV-1995 (first entry)

XX Partial ALK protein.

XX ALK protein; protein-tyrosine-kinase; fusion protein; NPM protein;  
 KM nucleolar phosphoprotein; anaplastic large cell lymphoma; t(2;  
 KM 5) translocation; diagnosis; gene therapy.  
 XX Homo sapiens.  
 OS  
 XX MO9515331-A1.  
 PN  
 XX 08-JUN-1995.  
 PD  
 XX  
 PF 05-DEC-1994; 94WO-US013947.  
 XX  
 PR 03-DEC-1993; 93US-00160861.  
 XX  
 PA (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 XX  
 PI Morris SW, Look AT;  
 XX  
 PT WPI; 1995-215226/28.  
 DR  
 DR N-PSDB; AAQ95541.  
 XX  
 PT Methods for detecting human t(2;5) lymphoma - for detection and diagnosis  
 PT of anaplastic large cell lymphoma(s).  
 XX  
 PS Claim 21; Page 40-42; 70pp; English.  
 XX  
 CC The translocation event that occurs in human t(2;5) lymphoma brings  
 CC sequences from the nucleolar phosphoprotein (NPM) gene on chromosome 5q35  
 CC to those from a protein-tyrosine-kinase (ALK) gene on chromosome 2q23.  
 CC Nucleic acids encoding a partial ALK protein (AAK8520) and the ALK/NPM  
 CC fusion protein (AAK78521) were isolated (AAQ95541, AAQ95542).  
 CC Identification of the NPM/ALK fusion allows lymphoma diagnosis, or  
 CC therapy through the use of antisense RNA or ribozymes. (Updated on 25-MAR  
 CC -2003 to correct PN field.)  
 CC  
 SQ Sequence 572 AA;  
 Query Match 100.0%; Score 52; DB 2; Length 572;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GVILMEIFSL 10  
 DB 423 GVILMEIFSL 432

RESULT 13  
 AAG20950  
 ID AAG20950 standard; protein; 675 AA.  
 XX  
 AC AAG20950;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #20941.  
 KM Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KM food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO200175067-A2.  
 PD  
 XX 11-OCT-2001.  
 PF 30-MAR-2001; 2001WO-US008631.  
 XX  
 PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX

PI Drmanac RT, Liu C, Tang YT;  
 XX  
 DR WPI; 2001-619362/73.  
 DR N-PSDB; AAB85137.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 PS Claim 20; SEQ ID NO 51309; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptides and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABB00010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pat\_sequences  
 XX  
 SQ Sequence 675 AA;  
 Query Match 100.0%; Score 52; DB 4; Length 675;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GVILMEIFSL 10  
 DB 371 GVILMEIFSL 380

RESULT 14  
 AAB73858  
 ID AAB73858 standard; protein; 680 AA.  
 XX  
 AC AAB73858;  
 XX  
 DT 15-MAY-2001 (first entry)  
 XX  
 DE Human NPM/ALK fusion protein.  
 KM Human; anaplastic lymphoma kinase; ALK; nucleophosmin; NPM;  
 KM chromosomal translocation; cancer; NPM/ALK; fusion.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6174674-B1.  
 PD  
 XX 16-JAN-2001.  
 PF 19-JUN-1998; 98US-00100089.  
 XX  
 PR 03-DEC-1993; 93US-00160861.  
 PR 12-OCT-1995; 95US-00542363.  
 XX  
 PA (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 XX  
 PI Morris SW, Look AT;  
 XX

DR WPI, 2001-243208/25.  
 DR N-PSDB; AAF76868.  
 XX  
 PT Detection of chromosomal rearrangement or translocations present in t(2;  
 PT 5) for cancer diagnosis comprises using and identifying anaplastic  
 PT lymphoma kinase and nucleophosmin genes in nucleic acid hybridization  
 PT techniques.  
 XX  
 PS Claim 1, Fig 2a; 87pp; English.  
 XX  
 CC The present sequence is given in a specification relating to a method for  
 CC detecting a chromosomal rearrangement involving a breakpoint in the  
 CC anaplastic lymphoma kinase (ALK) or nucleophosmin (NPM) gene. The method  
 CC is based on the identification and sequence determination of a novel  
 CC gene, ALK, which is fused to the gene encoding NPM in translocations  
 CC present in t(2;5) lymphoma cells. The method is useful for detecting a  
 CC t(2;5) chromosomal rearrangement, or a chromosomal rearrangement  
 CC involving the ALK gene or NPM gene. It is particularly useful in  
 CC determining if particular cells or tissues express ALK or NPM/ALK coding  
 CC sequences. The method is also useful in diagnostic assays to determine,  
 CC for example, if a mammal has cancer or a genetic predisposition to cancer  
 CC  
 SQ Sequence 680 AA;  
 Query Match 100.0%; Score 52; DB 4; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GVLLWEIFSL 10  
 DB 376 GVLLWEIFSL 385  
 RESULT 15  
 ABG95067  
 ID ABG95067 standard; protein; 680 AA.  
 AC  
 XX ABG95067;  
 DT 04-DEC-2002 (first entry)  
 XX  
 DE Human translocation (2; 5) (p23; q35) protein.  
 XX  
 KW Chromosome aberration; oncogenic fusion protein; cancer;  
 KW proliferative disease; cellular protein isoform; heat shock protein 90;  
 KW HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;  
 KW T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;  
 KW acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;  
 KW acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;  
 KW papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;  
 KW rhabdomyosarcoma; synovial sarcoma; viral infection.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200269900-A2.  
 XX  
 PD 12-SEP-2002.  
 XX  
 PF 01-MAR-2002; 2002WO-US006518.  
 XX  
 PR 01-MAR-2001; 2001US-0272751P.  
 XX  
 PA (CONF-) CONFORMA THERAPEUTICS CORP.  
 XX  
 PI Fritz LC, Burrows FJ;  
 XX  
 DR WPI; 2002-698710/75.  
 DR N-PSDB; AB873246.  
 XX  
 XX Treating genetically-defined disease associated with chromosomal  
 PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative  
 PT diseases, involves administering an inhibitor of heat shock protein 90.

PS Disclosure; Page 200-202; 389pp; English.  
 XX  
 CC The invention describes a method of treating genetically-defined disease  
 CC associated with chromosomal aberrations yielding oncogenic fusion  
 CC proteins (I), treating cancerous cells containing (I) in a heterogeneous  
 CC cell population, treating proliferative diseases associated with mutant  
 CC protein or cellular protein isoforms (II) dependent on heat shock protein  
 CC (HSP)-90, or selectively treating cells expressing (II) involving  
 CC administering HSP90-inhibitor. The method is useful for treating  
 CC genetically-defined disease with chromosomal aberration yielding  
 CC oncogenic fusion protein, treating cancerous cells containing fusion  
 CC protein in heterogeneous cell population, treating proliferative disease  
 CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or  
 CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.  
 CC p53), or selectively treating cells expressing mutant protein or cellular  
 CC protein isoform in a patient heterozygous for (II). The method is useful  
 CC for treating a disease e.g. haematopoietic disorder such as T or B cell  
 CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, AML, NHL and CMML,  
 CC or a disease characterised by a solid tumour such as papillary thyroid  
 CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and  
 CC synovial sarcoma. The method is also useful for treating viral  
 CC infections. This represents a protein encoded by the DNA sequence of a  
 CC chromosome aberration  
 XX  
 SQ Sequence 680 AA;  
 Query Match 100.0%; Score 52; DB 5; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GVLLWEIFSL 10  
 DB 376 GVLLWEIFSL 385

Search completed: November 30, 2004, 07:39:55  
 Job time : 79.4295 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: November 30, 2004, 07:19:22 ; Search time 376.923 Seconds  
(without alignments)  
9.410 Million cell updates/sec

Title: US-10-008-377A-2  
Perfect score: 52  
Sequence: 1 GVLIWEIFSL 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1575965 seqs, 354694765 residues

Total number of hits satisfying chosen parameters: 1575965

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	10	14	US-10-008-377-2
2	52	100.0	116	9	US-09-827-949-15
3	52	100.0	116	9	US-09-827-949-15
4	52	100.0	127	15	US-10-827-774-1411
5	52	100.0	290	16	US-10-763-418-28
6	52	100.0	291	16	US-10-763-418-28
7	52	100.0	294	16	US-10-763-418-17
8	52	100.0	299	15	US-10-334-143-202
9	52	100.0	317	9	US-09-939-754-5
10	52	100.0	317	9	US-09-939-832-5
11	52	100.0	317	9	US-09-939-833-5
12	52	100.0	348	9	US-09-982-610-18
13	52	100.0	366	16	US-10-664-421-26

14	52	100.0	367	9	US-09-939-754-9	Sequence 9, Appli
15	52	100.0	367	9	US-09-939-754-12	Sequence 12, Appli
16	52	100.0	367	9	US-09-939-832-9	Sequence 9, Appli
17	52	100.0	367	9	US-09-939-832-12	Sequence 12, Appli
18	52	100.0	367	9	US-09-939-833-9	Sequence 9, Appli
19	52	100.0	367	9	US-09-939-833-12	Sequence 12, Appli
20	52	100.0	370	16	US-10-664-421-170	Sequence 70, Appli
21	52	100.0	429	16	US-10-664-421-44	Sequence 44, Appli
22	52	100.0	567	9	US-10-327-949-4	Sequence 4, Appli
23	52	100.0	680	9	US-09-827-949-4	Sequence 4, Appli
24	52	100.0	1298	9	US-09-982-610-33	Sequence 33, Appli
25	52	100.0	1298	11	US-09-765-534B-2	Sequence 3, Appli
26	52	100.0	1298	14	US-10-262-538-12	Sequence 12, Appli
27	52	100.0	1298	14	US-10-165-133A-12	Sequence 12, Appli
28	52	100.0	1298	14	US-10-394-332A-67	Sequence 67, Appli
29	52	100.0	1298	16	US-10-325-046-4	Sequence 4, Appli
30	52	100.0	1298	17	US-10-774-802-2	Sequence 2, Appli
31	52	100.0	1298	17	US-10-663-176-32	Sequence 3, Appli
32	52	100.0	1311	11	US-09-765-534B-5	Sequence 5, Appli
33	52	100.0	1311	17	US-10-774-802-5	Sequence 5, Appli
34	52	100.0	1333	14	US-10-394-332A-65	Sequence 65, Appli
35	52	100.0	1338	14	US-10-060-065-23	Sequence 23, Appli
36	52	100.0	1338	14	US-10-059-585-44	Sequence 44, Appli
37	52	100.0	1338	14	US-10-021-660-84	Sequence 84, Appli
38	52	100.0	1338	14	US-10-090-183-4	Sequence 4, Appli
39	52	100.0	1338	14	US-10-116-275-129	Sequence 129, Appli
40	52	100.0	1338	15	US-10-211-462-215	Sequence 215, Appli
41	52	100.0	1338	15	US-10-621-116-26	Sequence 26, Appli
42	52	100.0	1339	14	US-10-165-133A-9	Sequence 9, Appli
43	52	100.0	1345	14	US-10-090-183-6	Sequence 6, Appli
44	52	100.0	1354	14	US-10-262-538-30	Sequence 30, Appli
45	52	100.0	1354	17	US-10-669-176-30	Sequence 30, Appli

## ALIGNMENTS

RESULT 1  
US-10-008-377-2  
; Sequence 2, Application US/10008377  
; Publication No. US20030157101A1  
; GENERAL INFORMATION:  
; APPLICANT: Gambacorti-Passerini, Carlo  
; TITLE OF INVENTION: Immunogenic ALK Peptides  
; FILE REFERENCE: 045922/241203  
; CURRENT APPLICATION NUMBER: US/10/008,377  
; CURRENT FILING DATE: 2001-11-15  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 10  
; TYPE: PPT  
; ORGANISM: Homo sapiens  
US-10-008-377-2

Query Match 100.0%; Score 52; DB 14; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.031;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLIWEIFSL 10  
DB 1 GVLIWEIFSL 10

RESULT 2  
US-09-827-949-15  
; Sequence 15, Application US/09827949  
; Patent No. US20010021505A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, Stephan W.  
; APPLICANT: Look, A. Thomas  
; TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof

FILE REFERENCE: 0656.0400004  
CURRENT APPLICATION NUMBER: US/09/827,949  
CURRENT FILING DATE: 2001-04-09  
PRIOR APPLICATION NUMBER: US 09/670,827  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 09/100,089  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: US 08/542,363  
PRIOR FILING DATE: 1995-10-12  
PRIOR APPLICATION NUMBER: US 08/160,861  
PRIOR FILING DATE: 1993-12-03  
NUMBER OF SEQ ID NOS: 43  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 15  
LENGTH: 116  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-827-949-15

Query Match 100.0%; Score 52; DB 9; Length 116;  
Best Local Similarity 100.0%; Pred. No. 0.32;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10  
|||||  
56 GVLMWEIFSL 65

RESULT 3  
US-09-827-949-19  
Sequence 19, Application US/09827949  
Patent No. US20010021505A1  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof  
FILE REFERENCE: 0656.0400004  
CURRENT APPLICATION NUMBER: US/09/827,949  
CURRENT FILING DATE: 2001-04-09  
PRIOR APPLICATION NUMBER: US 09/670,827  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 09/100,089  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: US 08/542,363  
PRIOR FILING DATE: 1995-10-12  
PRIOR APPLICATION NUMBER: US 08/160,861  
PRIOR FILING DATE: 1993-12-03  
NUMBER OF SEQ ID NOS: 43  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 19  
LENGTH: 116  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-827-949-19

Query Match 100.0%; Score 52; DB 9; Length 116;  
Best Local Similarity 100.0%; Pred. No. 0.32;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10  
|||||  
56 GVLMWEIFSL 65

RESULT 4  
US-10-276-774-1411  
Sequence 1411, Application US/10276774  
Publication No. US20040053245A1  
GENERAL INFORMATION:  
APPLICANT: Hyseq, Inc.  
TITLE OF INVENTION: Tany, Y, Tom et al  
FILE REFERENCE: 21272-030  
Nucleic Acids and Polypeptides

CURRENT APPLICATION NUMBER: US/10/276,774  
CURRENT FILING DATE: 2002-11-18  
PRIOR APPLICATION NUMBER: 09/560,875  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: 08/496,914  
PRIOR FILING DATE: 2000-02-03  
NUMBER OF SEQ ID NOS: 2700  
SOFTWARE: Custom  
SEQ ID NO 1411  
LENGTH: 127  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-276-774-1411

Query Match 100.0%; Score 52; DB 15; Length 127;  
Best Local Similarity 100.0%; Pred. No. 0.35;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10  
|||||  
108 GVLMWEIFSL 117

RESULT 5  
US-10-763-418-28  
Sequence 28, Application US/10763418  
Publication No. US20040185547A1  
GENERAL INFORMATION:  
APPLICANT: MOHAMMADI, MOOSA  
APPLICANT: SCHLESINGER, JOSEPH  
APPLICANT: HUBBARD, STEVAN R.  
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN  
FILE REFERENCE: 038602/0847  
CURRENT APPLICATION NUMBER: US/10/763,418  
CURRENT FILING DATE: 2004-01-26  
PRIOR APPLICATION NUMBER: US/09/664,526  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: 09/188,809  
PRIOR FILING DATE: 1998-11-09  
PRIOR APPLICATION NUMBER: 08/701,191  
PRIOR FILING DATE: 1996-08-21  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 28  
LENGTH: 290  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-763-418-28

Query Match 100.0%; Score 52; DB 16; Length 290;  
Best Local Similarity 100.0%; Pred. No. 0.77;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10  
|||||  
214 GVLMWEIFSL 223

RESULT 6  
US-10-377-268-28  
Sequence 28, Application US/10377268  
Publication No. US20040171062A1  
GENERAL INFORMATION:  
APPLICANT: HIRTH, KLAUS-PETER  
APPLICANT: MILBURN, MICHAEL VANCE  
TITLE OF INVENTION: METHOD FOR THE DESIGN OF MOLECULAR SCAFFOLDS AND LIGANDS  
FILE REFERENCE: 039363/0303  
CURRENT APPLICATION NUMBER: US/10/377,268  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: 60/437,929  
PRIOR FILING DATE: 2003-01-02  
PRIOR APPLICATION NUMBER: 60/360,651

PRIOR FILING DATE: 2002-02-28  
PRIOR APPLICATION NUMBER: 60/411,398  
PRIOR FILING DATE: 2002-09-16  
PRIOR APPLICATION NUMBER: 60/412,341  
PRIOR FILING DATE: 2002-09-20  
NUMBER OF SEQ ID NOS: 38  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 28  
LENGTH: 291  
TYPE: PRT  
ORGANISM: Unknown Organism  
FEATURE:  
OTHER INFORMATION: Description of Unknown Organism: KDR tyrosine kinase  
US-10-377-268-28

Query Match  
Best Local Similarity 100.0%; Score 52; DB 16; Length 291;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
|||  
Db 215 GVLWEIFSL 224

RESULT 7  
US-10-763-418-17  
Sequence 17, Application US/10763418  
Publication No. US20040185547A1  
GENERAL INFORMATION:  
APPLICANT: MOHAMMADI, MOOSA  
APPLICANT: SCHLESSINGER, JOSEPH  
APPLICANT: HUBBARD, STEVAN R.  
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN  
FILE REFERENCE: 038602/0847  
CURRENT FILING DATE: US/10/763,418  
PRIOR APPLICATION NUMBER: 2004-01-26  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: 09/186,809  
PRIOR FILING DATE: 1998-11-09  
PRIOR APPLICATION NUMBER: 08/701,191  
PRIOR FILING DATE: 1996-08-21  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 17  
LENGTH: 294  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-763-418-17

Query Match  
Best Local Similarity 100.0%; Score 52; DB 16; Length 294;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
|||  
Db 217 GVLWEIFSL 226

RESULT 8  
US-10-334-143-202  
Sequence 202, Application US/10334143  
Publication No. US20040009549A1  
GENERAL INFORMATION:  
APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH  
APPLICANT: SUDASANAM, SUCHA  
TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL  
FILE REFERENCE: 038602/1543  
CURRENT APPLICATION NUMBER: US/10/334,143  
CURRENT FILING DATE: 2002-12-31

PRIOR APPLICATION NUMBER: 60/343,169  
PRIOR FILING DATE: 2001-12-31  
NUMBER OF SEQ ID NOS: 207  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 202  
LENGTH: 299  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: MOD\_RES  
LOCATION: (190)  
OTHER INFORMATION: Variable amino acid  
US-10-334-143-202

Query Match  
Best Local Similarity 100.0%; Score 52; DB 15; Length 299;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
|||  
Db 223 GVLWEIFSL 232

RESULT 9  
US-09-939-754-5  
Sequence 5, Application US/09939754  
Patent No. US20020051965A1  
GENERAL INFORMATION:  
APPLICANT: MCTIGUE, MICHELE A.  
APPLICANT: WICKERSHAM, JOHN A.  
APPLICANT: PINKO, CHRIS  
APPLICANT: SHOWALTER, RICHARD  
APPLICANT: PARAST, CAMRAN V.  
APPLICANT: TEMPICZYK-RUSSEL, ANNA  
APPLICANT: GERING, MICHAEL R.  
APPLICANT: WROCKZKOWSKI, BARBARA  
APPLICANT: KAN, CHEN-CHEN  
APPLICANT: VILLAFRANCA, J. ERNEST  
APPLICANT: APPELT, KRYSZTOF  
TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND  
METHODS OF USE  
FILE REFERENCE: 0125-0016US  
CURRENT FILING DATE: US/09/939,754  
PRIOR APPLICATION NUMBER: 09/390,326  
PRIOR FILING DATE: 1993-09-07  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO: 5  
LENGTH: 317  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-939-754-5

Query Match  
Best Local Similarity 100.0%; Score 52; DB 9; Length 317;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
|||  
Db 238 GVLWEIFSL 247

RESULT 10  
US-09-939-832-5  
Sequence 5, Application US/09939832  
Patent No. US20020127538A1  
GENERAL INFORMATION:  
APPLICANT: MCTIGUE, MICHELE A.  
APPLICANT: WICKERSHAM, JOHN A.  
APPLICANT: PINKO, CHRIS  
APPLICANT: SHOWALTER, RICHARD  
APPLICANT: PARAST, CAMRAN V.

APPLICANT: TEMPCZYK-RUSSEL, ANNA  
APPLICANT: GEHRING, MICHAEL R.  
APPLICANT: MROCKOWSKI, BARBARA  
APPLICANT: KAN, CHEN-CHEN  
APPLICANT: VILLAFRANCA, J. ERNEST  
APPLICANT: APPELT, KRZYSZTOF  
TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND  
FILE REFERENCE: 0125-0016US  
CURRENT APPLICATION NUMBER: US/09/939,832  
CURRENT FILING DATE: 2001-08-28  
PRIOR APPLICATION NUMBER: 09/390,326  
PRIOR FILING DATE: 1999-09-07  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 317  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-939-832-5

Query Match 100.0%; Score 52; DB 9; Length 317;  
Best Local Similarity 100.0%; Pred. No. 0.84;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVLWEIFSL 10  
Db 238 GVLWEIFSL 247

RESULT 11  
US-09-939-833-5  
Sequence 5, Application US/0939833  
Patent No. US2002016441A1  
GENERAL INFORMATION:  
APPLICANT: MCTIGUE, MICHELE A.  
APPLICANT: WICKERSHAM, JOHN A.  
APPLICANT: PINKO, CHRIS  
APPLICANT: SHOWALTER, RICHARD  
APPLICANT: PARAST, CAMRAN V.  
APPLICANT: TEMPCZYK-RUSSEL, ANNA  
APPLICANT: GEHRING, MICHAEL R.  
APPLICANT: MROCKOWSKI, BARBARA  
APPLICANT: KAN, CHEN-CHEN  
APPLICANT: VILLAFRANCA, J. ERNEST  
APPLICANT: APPELT, KRZYSZTOF  
TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND  
FILE REFERENCE: 0125-0016US  
CURRENT APPLICATION NUMBER: US/09/939,833  
CURRENT FILING DATE: 2001-08-28  
PRIOR APPLICATION NUMBER: US/09/390,326  
PRIOR FILING DATE: 1999-09-07  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 317  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-939-833-5

Query Match 100.0%; Score 52; DB 9; Length 317;  
Best Local Similarity 100.0%; Pred. No. 0.84;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVLWEIFSL 10  
Db 238 GVLWEIFSL 247

RESULT 12  
US-09-982-610-18  
Sequence 18, Application US/09982610

Patent No. US2002016420A1

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.

Bennett, Brian D.

Goeddel, David

Lee, James M.

Matthews, William

Tsai, Siao Ping

Wood, William I.

TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/982,610

FILING DATE: 17-Oct-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/446,648

FILING DATE: 1996-MAY-23

APPLICATION NUMBER: 08/222,616

FILING DATE: 04-APR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Lee, Wendy M.

REGISTRATION NUMBER: 40,378

REFERENCE/DOCKET NUMBER: P0821P3PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-1994

TELEFAX: 415/852-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:

LENGTH: 348 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

SEQUENCE DESCRIPTION: SEQ ID NO: 18:

US-09-982-610-18

Query Match 100.0%; Score 52; DB 9; Length 348;  
Best Local Similarity 100.0%; Pred. No. 0.91;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVLWEIFSL 10  
Db 151 GVLWEIFSL 160

RESULT 13  
US-10-664-421-26

Sequence 26, Application US/10664421

Publication No. US2004014286A1

GENERAL INFORMATION:

APPLICANT: BREWER, RYAN

APPLICANT: IERHIM, PRABHA

APPLICANT: KUMAR, ABHINAV

APPLICANT: MANDIVAN, VALSAN V.

APPLICANT: MILBURN, MICHAEL V.

TITLE OF INVENTION: CRYSTAL STRUCTURE OF PIM-1 KINASE

FILE REFERENCE: 039363/0703

CURRENT APPLICATION NUMBER: US/10/664,421

PRIOR FILING DATE: 2003-09-16

PRIOR APPLICATION NUMBER: 60/412,341

PRIOR FILING DATE: 2002-09-20

PRIOR APPLICATION NUMBER: 60/411,398  
PRIOR FILING DATE: 2002-09-16  
NUMBER OF SEQ ID NOS: 169  
SOFTWARE: PatentIn Ver. 3.2  
SEQ ID NO 26  
LENGTH: 366  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-664-421-26

Query Match 100.0%; Score 52; DB 16; Length 366;  
Best Local Similarity 100.0%; Pred. No. 0.96;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMEIFSL 10  
|||||  
DB 287 GVLMEIFSL 297

RESULT 14  
US-09-939-754-9  
Sequence 9, Application US/09939754  
Patent No. US20020051965A1  
GENERAL INFORMATION:  
APPLICANT: MCTIGUE, MICHELE A.  
APPLICANT: WICKERSHAM, JOHN A.  
APPLICANT: PINKO, CHRIS  
APPLICANT: SHOWALTER, RICHARD  
APPLICANT: PARAST, CAMRAN V.  
APPLICANT: TEMPCZYK-RUSSEL, ANNA  
APPLICANT: GEHRING, MICHAEL R.  
APPLICANT: MROCKZOWSKI, BARBARA  
APPLICANT: KAN, CHEN-CHEN  
APPLICANT: VILAFRANCA, J. ERNEST  
APPLICANT: APPELT, KRZYSZTOF  
TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND  
FILE REFERENCE: 0125-0016US  
CURRENT APPLICATION NUMBER: US/09/939,754  
CURRENT FILING DATE: 2001-08-28  
PRIOR APPLICATION NUMBER: 09/390,326  
PRIOR FILING DATE: 1999-09-07  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 9  
LENGTH: 367  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-939-754-9

Query Match 100.0%; Score 52; DB 9; Length 367;  
Best Local Similarity 100.0%; Pred. No. 0.96;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMEIFSL 10  
|||||  
DB 288 GVLMEIFSL 297

RESULT 15  
US-09-939-754-12  
Sequence 12, Application US/09939754  
Patent No. US20020051965A1  
GENERAL INFORMATION:  
APPLICANT: MCTIGUE, MICHELE A.  
APPLICANT: WICKERSHAM, JOHN A.  
APPLICANT: PINKO, CHRIS  
APPLICANT: SHOWALTER, RICHARD  
APPLICANT: PARAST, CAMRAN V.  
APPLICANT: TEMPCZYK-RUSSEL, ANNA  
APPLICANT: GEHRING, MICHAEL R.  
APPLICANT: MROCKZOWSKI, BARBARA  
APPLICANT: KAN, CHEN-CHEN

APPLICANT: VILAFRANCA, J. ERNEST  
APPLICANT: APPELT, KRZYSZTOF  
TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND  
FILE REFERENCE: 0125-0016US  
CURRENT APPLICATION NUMBER: US/09/939,754  
CURRENT FILING DATE: 2001-08-28  
PRIOR APPLICATION NUMBER: 09/390,326  
PRIOR FILING DATE: 1999-09-07  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 12  
LENGTH: 367  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-939-754-12

Query Match 100.0%; Score 52; DB 9; Length 367;  
Best Local Similarity 100.0%; Pred. No. 0.96;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMEIFSL 10  
|||||  
DB 288 GVLMEIFSL 297

Search completed: November 30, 2004, 08:06:35  
Job time : 377.923 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 ; Search time 25.1282 Seconds  
(without alignments)  
26.392 Million cell updates/sec

Title: US-10-008-377A-2

Perfect score: 52

Sequence: 1 GVLMWEIFSL 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database:

1: Issued Patents AA.\*  
2: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	100.0	116	1	US-08-542-363-15
2	52	100.0	116	1	US-08-542-363-15
3	52	100.0	116	3	US-09-100-089-15
4	52	100.0	116	3	US-09-100-089-15
5	52	100.0	116	4	US-09-670-827-15
6	52	100.0	116	4	US-09-670-827-15
7	52	100.0	116	4	US-09-827-949-15
8	52	100.0	116	4	US-09-827-949-15
9	52	100.0	290	2	US-08-701-191A-28
10	52	100.0	290	2	US-08-701-191A-28
11	52	100.0	294	2	US-08-664-526-28
12	52	100.0	294	2	US-08-701-191A-17
13	52	100.0	317	3	US-09-390-326-5
14	52	100.0	317	3	US-09-390-326-5
15	52	100.0	317	4	US-09-939-833-5
16	52	100.0	317	4	US-09-939-833-5
17	52	100.0	317	4	US-09-939-832-5
18	52	100.0	348	1	US-08-222-516-18
19	52	100.0	348	1	US-08-446-648-18
20	52	100.0	348	5	US-09-982-610-18
21	52	100.0	367	3	PCT-US95-04228-18
22	52	100.0	367	3	US-09-390-326-9
23	52	100.0	367	3	US-09-390-326-12
24	52	100.0	367	4	US-09-939-833-9
25	52	100.0	367	4	US-09-939-833-12
26	52	100.0	367	4	US-09-506-906-9
27	52	100.0	367	4	US-09-506-906-12
					Sequence 9, Appl1

28	52	100.0	367	4	US-09-939-832-12	Sequence 12, Appl1
29	52	100.0	525	1	US-08-160-861-4	Sequence 4, Appl1
30	52	100.0	572	1	US-08-160-861-3	Sequence 3, Appl1
31	52	100.0	680	1	US-08-542-363-4	Sequence 4, Appl1
32	52	100.0	680	4	US-09-100-089-4	Sequence 4, Appl1
33	52	100.0	680	4	US-09-670-827-4	Sequence 4, Appl1
34	52	100.0	680	4	US-09-827-949-4	Sequence 4, Appl1
35	52	100.0	1298	1	US-08-222-616-33	Sequence 33, Appl1
36	52	100.0	1298	1	US-08-340-011-2	Sequence 2, Appl1
37	52	100.0	1298	3	US-08-901-710-2	Sequence 2, Appl1
38	52	100.0	1298	3	US-08-446-648-33	Sequence 33, Appl1
39	52	100.0	1298	4	US-09-982-610-33	Sequence 33, Appl1
40	52	100.0	1298	5	PCT-US95-04228-33	Sequence 33, Appl1
41	52	100.0	1311	1	US-08-340-011-5	Sequence 5, Appl1
42	52	100.0	1311	3	US-08-901-710-5	Sequence 5, Appl1
43	52	100.0	1338	3	US-08-750-141A-3	Sequence 3, Appl1
44	52	100.0	1338	4	US-09-119-014D-6	Sequence 6, Appl1
45	52	100.0	1356	1	US-08-810-116-8	Sequence 8, Appl1

#### ALIGNMENTS

RESULT 1  
US-08-542-363-15  
Sequence 15, Application US/08542363  
Parent No. 5770421  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
APPLICANT: Look, A. Thomas  
TITLE OF INVENTION: Alk Protein Tyrosine Kinase/Receptor and  
TITLE OF INVENTION: Ligands Thereof  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSER: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/542,363  
FILING DATE: 12-OCT-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400001/SLF/GKT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 116 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-542-363-15

Query Match 100.0%; Score 52; DB 1; Length 116;  
Best Local Similarity 100.0%; Pred. No. 0.052;  
Matches 10; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;  
CY 1 GVLMWEIFSL 10  
DB 56 GVLMWEIFSL 65

RESULT 2  
US-08-542-363-19  
Sequence 19, Application US/08542363  
Patent No. 5770421  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
APPLICANT: Look, A. Thomas  
TITLE OF INVENTION: Alk Protein Tyrosine Kinase/Receptor and  
TITLE OF INVENTION: Ligands Thereof  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/542,363  
FILING DATE: 12-OCT-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400001/SLE/GKT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2540  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 116 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-542-363-19  
Query Match 100.0%; Score 52; DB 1; Length 116;  
Best Local Similarity 100.0%; Pred. No. 0.052;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
Db 56 GVLWEIFSL 65

RESULT 3  
US-09-100-089-15  
Sequence 15, Application US/09100089  
Patent No. 6174674  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
APPLICANT: Look, A. Thomas  
TITLE OF INVENTION: Alk Protein Tyrosine Kinase/Receptor and  
TITLE OF INVENTION: Ligands Thereof  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/100,089  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2540  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 116 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-100-089-15  
Query Match 100.0%; Score 52; DB 3; Length 116;  
Best Local Similarity 100.0%; Pred. No. 0.052;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
Db 56 GVLWEIFSL 65

RESULT 4  
US-09-100-089-19  
Sequence 19, Application US/09100089  
Patent No. 6174674  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
APPLICANT: Look, A. Thomas  
TITLE OF INVENTION: Alk Protein Tyrosine Kinase/Receptor and  
TITLE OF INVENTION: Ligands Thereof  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/100,089  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400002  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 116 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-100-089-19

Query Match 100.0%; Score 52; DB 3; Length 116;  
Best Local Similarity 100.0%; Pred. No. 0.052;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
DB 56 GVLWEIFSL 65

## RESULT 5

US-09-670-827-15  
Sequence 15, Application US/09670827  
Patent No. 6451997  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
Look, A. Thomas  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
Ligands Thereof  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/670,827  
FILING DATE: 28-SEP-2000  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995  
APPLICATION NUMBER: US 09/100,089  
FILING DATE: 19-JUN-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 116 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 15:  
US-09-670-827-15

Query Match 100.0%; Score 52; DB 4; Length 116;  
Best Local Similarity 100.0%; Pred. No. 0.052;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
DB 56 GVLWEIFSL 65

## RESULT 6

US-09-670-827-19  
Sequence 19, Application US/09670827  
Patent No. 6451997  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
Look, A. Thomas  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
Ligands Thereof  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/670,827  
FILING DATE: 28-SEP-2000  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-DEC-1993  
APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-OCT-1995  
APPLICATION NUMBER: US 09/100,089  
FILING DATE: 19-JUN-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0400003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 116 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-09-670-827-19

Query Match 100.0%; Score 52; DB 4; Length 116;  
Best Local Similarity 100.0%; Pred. No. 0.052;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLWEIFSL 10  
DB 56 GVLWEIFSL 65

## RESULT 7

US-09-827-949-15  
Sequence 15, Application US/09627949  
Patent No. 6696548  
GENERAL INFORMATION:  
APPLICANT: Morris, Stephan W.  
Look, A. Thomas  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof

```
FILE REFERENCE: 0656.0400004
CURRENT APPLICATION NUMBER: US/09/827,949
CURRENT FILING DATE: 2001-04-09
PRIOR APPLICATION NUMBER: US 09/670,827
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 09/100,089
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: US 08/542,363
PRIOR FILING DATE: 1995-10-12
PRIOR APPLICATION NUMBER: US 08/160,861
PRIOR FILING DATE: 1993-12-03
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn version 3.0
SEQ ID NO 15
LENGTH: 116
TYPE: PRT
ORGANISM: Homo sapiens
US-09-827-949-15

Query Match      100.0%; Score 52; DB 4; Length 116;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

CY 1 GVLWEIIFSL 10  
|||  
56 GVLWEIIFSL 65

Db

```
RESULT 8
US-09-827-949-19
Sequence 19, Application US/09827949
Patent No. 6696548
GENERAL INFORMATION:
APPLICANT: Morris, Stephan W.
TITLE OF INVENTION: ALK Protease Tyrosine Kinase/Receptor and Ligands Thereof
FILE REFERENCE: 0656.0400004
CURRENT APPLICATION NUMBER: US/09/827,949
CURRENT FILING DATE: 2001-04-09
PRIOR APPLICATION NUMBER: US 09/670,827
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 09/100,089
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: US 08/542,363
PRIOR FILING DATE: 1995-10-12
PRIOR APPLICATION NUMBER: US 08/160,861
PRIOR FILING DATE: 1993-12-03
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn version 3.0
SEQ ID NO 19
LENGTH: 116
TYPE: PRT
ORGANISM: Homo sapiens
US-09-827-949-19

Query Match      100.0%; Score 52; DB 4; Length 116;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

CY 1 GVLWEIIFSL 10  
|||  
56 GVLWEIIFSL 65

Db

```
RESULT 9
US-08-701-191A-28
Sequence 28, Application US/08701191A
Patent No. 5942428
GENERAL INFORMATION:
APPLICANT: Moosa Mohammadi, Joseph Schllessinger,
APPLICANT: and Stevan R. Hubbard
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
TITLE OF INVENTION: OF NON-INSULIN RECEPTOR TYROSINE KINASE
```

```
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/701,191A
FILING DATE: August 21, 1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 227/088
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 955-0440
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 290 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-701-191A-28

Query Match      100.0%; Score 52; DB 2; Length 290;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

CY 1 GVLWEIIFSL 10  
|||  
214 GVLWEIIFSL 223

Db

```
RESULT 10
US-09-664-526-28
Sequence 28, Application US/09664526
Patent No. 6682821
GENERAL INFORMATION:
APPLICANT: MOHAMMADI, MOOSA
APPLICANT: SCHLESSINGER, JOSEPH
APPLICANT: HUBBARD, STEVAN R.
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
TITLE OF INVENTION: RECEPTOR TYROSINE KINASES
FILE REFERENCE: 038602/0847
CURRENT APPLICATION NUMBER: US/09/664,526
CURRENT FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/188,809
PRIOR FILING DATE: 1998-11-09
PRIOR APPLICATION NUMBER: 08/701,191
PRIOR FILING DATE: 1996-08-21
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 28
LENGTH: 290
TYPE: PRT
ORGANISM: Homo sapiens
US-09-664-526-28
```

Query Match 100.0%; Score 52; DB 4; Length 290;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10  
DB 214 GVLMWEIFSL 223

## RESULT 11

US-08-701-191A-17  
Sequence 17, Application US/08701191A  
Patent No. 5942428  
GENERAL INFORMATION:  
APPLICANT: Moosa Mohammadi, Joseph Schlessinger,  
and Stevan R. Hubbard  
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN  
OF NON-INSULIN RECEPTOR TYROSINE KINASE  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: Storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FastSeq for Windows 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/701.191A  
FILING DATE: August 21, 1996  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Marbury, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 227/098  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 294 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-701-191A-17

Query Match 100.0%; Score 52; DB 2; Length 294;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10  
DB 217 GVLMWEIFSL 226

## RESULT 12

US-09-664-526-17  
Sequence 17, Application US/09664526  
Patent No. 6682321  
GENERAL INFORMATION:  
APPLICANT: MOHAMMADI, MOOSA  
APPLICANT: SCHLESSINGER, JOSEPH

APPLICANT: HUBBARD, STEVAN R.  
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN  
RECEPTOR TYROSINE KINASES  
FILE REFERENCE: 038602/0847  
CURRENT APPLICATION NUMBER: US/09/664,526  
CURRENT FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: 09/188,809  
PRIOR FILING DATE: 1998-11-09  
PRIOR APPLICATION NUMBER: 08/701,191  
PRIOR FILING DATE: 1996-08-21  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 17  
LENGTH: 294  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-664-526-17

Query Match 100.0%; Score 52; DB 4; Length 294;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10  
DB 217 GVLMWEIFSL 226

## RESULT 13

US-09-390-326-5  
Sequence 5, Application US/09390326  
Patent No. 6316603  
GENERAL INFORMATION:  
APPLICANT: MCTIGUE, MICHELE A.  
APPLICANT: WICKERSHAM, JOHN A.  
APPLICANT: PINKO, CHRIS  
APPLICANT: SHWALTER, RICHARD  
APPLICANT: PARAST, CAMRAN V.  
APPLICANT: TEMPYAK-RUSSEL, ANNA  
APPLICANT: GEHRING, MICHAEL R.  
APPLICANT: MROCKZOMSKI, BARBARA  
APPLICANT: KAN, CHEN-CHEN  
APPLICANT: VILLAFRANCA, J. ERNEST  
APPLICANT: APPELT, KRYSZTOF  
TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND  
METHODS OF USE  
FILE REFERENCE: 0125-0016US  
CURRENT APPLICATION NUMBER: US/09/390,326  
CURRENT FILING DATE: 1999-09-07  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 317  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-390-326-5

Query Match 100.0%; Score 52; DB 3; Length 317;  
Best Local Similarity 100.0%; Pred. No. 0.15;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLMWEIFSL 10  
DB 238 GVLMWEIFSL 247

## RESULT 14

US-09-939-833-5  
Sequence 5, Application US/09939833  
Patent No. 6753416  
GENERAL INFORMATION:  
APPLICANT: MCTIGUE, MICHELE A.  
APPLICANT: WICKERSHAM, JOHN A.  
APPLICANT: PINKO, CHRIS

Job time : 25.1262 secs

APPLICANT: SHOWALTER, RICHARD  
APPLICANT: PARAST, CAMRAN V.  
APPLICANT: TEMPICYZK-RUSSEL, ANNA  
APPLICANT: GERING, MICHAEL R.  
APPLICANT: MROCKZKOWSKI, BARBARA  
APPLICANT: KAN, CHEN-CHEN  
APPLICANT: VILLAFRANCA, J. ERNEST  
APPLICANT: APPELT, KRZYSZTOF  
TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND  
FILE REFERENCE: 0125-0016US  
CURRENT APPLICATION NUMBER: US/09/939,833  
PRIOR FILING DATE: 2001-08-28  
PRIOR APPLICATION NUMBER: US/09/390,326  
PRIOR FILING DATE: 1999-09-07  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO: 5  
LENGTH: 317  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-939-833-5

Query Match 100.0%; Score 52; DB 4; Length 317;  
Best Local Similarity 100.0%; Pred. No. 0.15;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLLWEIFSL 10  
|||  
Db 238 GVLLWEIFSL 247

RESULT 15  
US-09-506-906-5  
Sequence 5, Application US/09506906  
Patent No. 6784285  
GENERAL INFORMATION:  
APPLICANT: MCTIGUE, MICHELE A.  
APPLICANT: WICKERSHAM, JOHN A.  
APPLICANT: PINKO, CHRIS  
APPLICANT: SHOWALTER, RICHARD  
APPLICANT: PARAST, CAMRAN V.  
APPLICANT: TEMPICYZK-RUSSEL, ANNA  
APPLICANT: GERING, MICHAEL R.  
APPLICANT: MROCKZKOWSKI, BARBARA  
APPLICANT: KAN, CHEN-CHEN  
APPLICANT: VILLAFRANCA, J. ERNEST  
APPLICANT: APPELT, KRZYSZTOF  
TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND  
FILE REFERENCE: 0125-0016US  
CURRENT APPLICATION NUMBER: US/09/506,906  
CURRENT FILING DATE: 2000-02-18  
PRIOR APPLICATION NUMBER: US/09/390,326  
PRIOR FILING DATE: 1999-09-07  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO: 5  
LENGTH: 317  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-506-906-5

Query Match 100.0%; Score 52; DB 4; Length 317;  
Best Local Similarity 100.0%; Pred. No. 0.15;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVLLWEIFSL 10  
|||  
Db 238 GVLLWEIFSL 247

Search completed: November 30, 2004, 07:19:15

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 ; Search time 20 Seconds

(without alignments)  
48.108 Million cell updates/sec

Title: US-10-008-377A-6

Sequence: 1 ALPIEYGPLV 10

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR79:\*  
2: PIR1:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	680	4	138491
2	49	94.2	1621	2	T30200
3	41	78.8	116	2	AG2816
4	41	78.8	116	2	H97594
5	38	73.1	163	2	D82167
6	38	73.1	401	2	H83179
7	37	71.2	201	2	S51424
8	37	71.2	1083	2	H86921
9	36	69.2	203	2	S59279
10	36	69.2	233	2	H83544
11	36	69.2	502	2	H86264
12	36	69.2	502	2	TS2173
13	36	69.2	1078	2	PC4198
14	35	67.3	76	2	I45524
15	35	67.3	216	2	H83166
16	35	67.3	235	2	H83951
17	35	67.3	274	2	S74680
18	35	67.3	344	2	AE3642
19	35	67.3	402	2	T48729
20	35	67.3	507	2	T08337
21	35	67.3	526	2	T04566
22	35	67.3	814	2	T00740
23	35	67.3	897	2	S76958
24	35	67.3	1207	2	H87475
25	35	67.3	7576	2	T17428
26	34	65.4	292	2	B17659
27	34	65.4	57	1	T18PVA
28	34	65.4	61	2	AF2219
29	34	65.4	77	2	F88165

30	34	65.4	123	2	AC3470	hypothetical prote
31	34	65.4	127	2	F88801	protein CO492.4 [1]
32	34	65.4	127	2	T21640	hypothetical prote
33	34	65.4	127	2	T24885	hypothetical prote
34	34	65.4	127	2	F88146	protein C34F11.6 [1]
35	34	65.4	127	2	A88139	protein ZK546.6 [1]
36	34	65.4	127	2	C88164	protein K05F1.7 [1]
37	34	65.4	127	2	H88146	protein C34F11.4 [1]
38	34	65.4	127	2	G88686	protein msp-19 [1]
39	34	65.4	127	2	A88165	protein ZK1248.6 [1]
40	34	65.4	127	2	B88689	protein msp-65 [1]
41	34	65.4	127	2	F88138	protein msp-31 [1]
42	34	65.4	127	2	G88145	protein F58A6.8 [1]
43	34	65.4	127	2	D88164	protein msp-142 [1]
44	34	65.4	127	2	H88688	protein msp-59 [1]
45	34	65.4	127	2	C88689	protein msp-51 [1]

#### ALIGNMENTS

##### RESULT 1

138491 nucleophosmin/anaplastic lymphoma kinase mutant fusion protein - human

C/Species: Homo sapiens (man)

C/Date: 15-Feb-1996 #sequence\_revision 15-Feb-1996 #text\_change 20-Apr-2000

C/Accession: 138491

R/Morris, S.W.; Kirstein, M.N.; Valentine, M.B.; Dittmer, K.G.; Shapiro, D.N.; Saltman, Science 263, 1281-1284, 1994

A/Title: Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma

A/Reference number: A53182; MIMD:94167588; PMID:8122112

A/Accession: 138491

A/Molecule type: mRNA

A/Residues: 1-680 <MOR>

A/Cross-references: EMBL:U04946; NID:G609341; PIDN:AA58698.1; PID:G609342

C/Comment: This sequence is the chimeric product of a translocation mutation.

C/Genetics:

A/Map position: 5/2p23-2p23

C/Keywords: fusion protein

Query Match	100.0%	Score 52	DB 4	Length 680
Best Local Similarity	100.0%	Pred. No. 0.058		
Matches 10	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1	ALPIEYGPLV 10		
Db	456	ALPIEYGPLV 465		

##### RESULT 2

T30200 protein-tyrosine kinase (EC 2.7.1.112) alk - mouse

N/Alternate names: anaplastic lymphoma kinase

C/Species: Mus musculus (house mouse)

C/Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T30200

R/Ishihara, T.; Fujimoto, J.; Wen, D.; Cupples, R.; Bucay, N.; Arakawa, T.; Mori, S.; Ra

Oncogene 14, 439-449, 1997

A/Title: Molecular characterization of ALK, a receptor tyrosine kinase expressed specif

A/Reference number: Z20774; MIMD:97178863; PMID:9035841

A/Accession: T30200

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-1621 <TMA>

A/Cross-references: UNIPROT:P9793; EMBL:DB3002; NID:q186406; PIDN:BA11673.1; PID:918

A/Experimental source: brain and testis

C/Genetics:

A/Gene: alk

C/Function:

A/Description: may play an important role in development of the brain

C/Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase

Query Match 94.2%; Score 49; DB 2; Length 1621;  
 Best Local Similarity 90.0%; Pred. No. 0.58;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIEYGPL 10  
 |||||  
 Db 1400 ALPIEYGPLV 1400

## RESULT 3

AG2816  
 conserved hypothetical protein Atu1953 [imported] - Agrobacterium tumefaciens (strain C5)  
 C.Species: Agrobacterium tumefaciens  
 C.Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
 C.Accession: AG2816  
 R.Wood, D.W., Setubal, J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E., Chen, Y., Woo, I.  
 erage, G., Giller, W., Grant, C., Genthner, D., Kutyavin, T., Levy, R., Li, M., McClellan,  
 : Karp, P., Romero, P., Zhang, S.  
 Science 294, 2317-2323, 2001  
 A.Authors: Yoo, H.; Tao, Y.; Biddle, P.; Ung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
 ster, R.W.  
 A.Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A.Reference number: AB2577; MIMD:21608550; PMID:11743193  
 A.Accession: AG2816  
 A.Status: preliminary  
 A.Molecule type: DNA  
 A.Residues: 1-116 <KUR>  
 A.Cross-references: UNIPROT:Q8UE11; GB:AE008688; PIND:AA12949.1; PID:G17740407; GSPDB:G  
 A.Experimental source: strain C58 (Dugout)  
 C.Genetics:  
 A.Gene: Atu1953  
 A.Map position: circular chromosome

QY 2 LPPIEYGPL 9  
 |||||  
 Db 84 LPPIEYGPLV 91

Query Match 78.8%; Score 41; DB 2; Length 116;  
 Best Local Similarity 87.5%; Pred. No. 1.1;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPPIEYGPL 9  
 |||||  
 Db 84 LPPIEYGPLV 91

## RESULT 4

H97594  
 hypothetical protein AGR\_C\_3564 [imported] - Agrobacterium tumefaciens (strain C58, Cere  
 C.Species: Agrobacterium tumefaciens  
 C.Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
 C.Accession: H97594  
 R.Goodner, B., Hinkle, G., Gattung, S., Miller, N., Blanchard, M., Qurollo, B., Goldman,  
 A., Liu, F., Mollam, C., Allinger, M., Dougherty, D., Scott, C., Lappas, C., Markelz, B.,  
 Science 294, 2323-2328, 2001  
 A.Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
 A.Reference number: A97359; MIMD:21608551; PMID:11743194  
 A.Accession: H97594  
 A.Status: preliminary  
 A.Molecule type: DNA  
 A.Residues: 1-116 <KUR>  
 A.Cross-references: UNIPROT:Q8UE11; GB:AE007869; PIND:AAK6773.1; PID:G15157076; GSPDB:G  
 C.Genetics:  
 A.Gene: AGR\_C\_3564  
 A.Map position: circular chromosome

QY 2 LPPIEYGPL 9  
 |||||  
 Db 84 LPPIEYGPLV 91

Query Match 78.8%; Score 41; DB 2; Length 116;  
 Best Local Similarity 87.5%; Pred. No. 1.1;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPPIEYGPL 9  
 |||||  
 Db 84 LPPIEYGPLV 91

## RESULT 5

DB2167

conserved hypothetical protein VC1703 [imported] - Vibrio cholerae (strain N16961 serogr  
 C.Species: Vibrio cholerae  
 C.Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
 C.Accession: DB2167  
 R.Heldberg, U.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L., Dodson, R.J.,  
 Chardson, D., Ermolaeva, M.D., Vamathevan, J., Bass, S., Qin, H., Dragoti, I., Sellers, F  
 l, R.R., Mekalanos, J.C., Venter, J.C., Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A.Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
 A.Reference number: AB2035; MIMD:20406833; PMID:10952301  
 A.Accession: DB2167  
 A.Status: preliminary  
 A.Molecule type: DNA  
 A.Residues: 1-163 <HNT>  
 A.Cross-references: UNIPROT:Q9KRD9; GB:AE004248; GB:AE003852; NID:g9656216; PIND:AAE948t  
 A.Experimental source: serogroup O1, strain N16961; biotype El Tor  
 C.Genetics:  
 A.Gene: VC1703  
 A.Map position: 1

QY 2 LPPIEYGPL 9  
 |||||  
 Db 76 LPPIEYGPL 83

Query Match 73.1%; Score 38; DB 2; Length 163;  
 Best Local Similarity 75.0%; Pred. No. 6.3;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPPIEYGPL 9  
 |||||  
 Db 76 LPPIEYGPL 83

## RESULT 6

FB3179  
 hypothetical protein PA3734 [imported] - Pseudomonas aeruginosa (strain PA01)  
 C.Species: Pseudomonas aeruginosa  
 C.Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
 C.Accession: FB3179  
 R.Stover, C.K., Pham, X.Q., Erwin, A.L., Mizoguchi, S.D., Warren, P., Hickey, M.J., B  
 adman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Tarrig, K., Lim,  
 ., Lory, S., Olson, M.V.  
 Nature 406, 959-964, 2000  
 A.Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path  
 A.Reference number: AB2950; MIMD:20437337; PMID:10984043  
 A.Accession: FB3179  
 A.Status: preliminary  
 A.Molecule type: DNA  
 A.Residues: 1-401 <STO>  
 A.Cross-references: UNIPROT:Q9HXQ9; GB:AE004792; GB:AE004091; NID:g9949895; PIND:AA6071  
 A.Experimental source: strain PA01  
 C.Genetics:  
 A.Gene: PA3734

QY 1 ALPIEYGPL 9  
 |||||  
 Db 278 ALPIEYGPL 286

Query Match 73.1%; Score 38; DB 2; Length 401;  
 Best Local Similarity 88.9%; Pred. No. 17;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALPIEYGPL 9  
 |||||  
 Db 278 ALPIEYGPL 286

## RESULT 7

SS1424  
 hypothetical protein YLR179c - yeast (Saccharomyces cerevisiae)  
 N.Alternate names: hypothetical protein L9470.20  
 C.Species: Saccharomyces cerevisiae  
 C.Date: 23-Feb-1995 #sequence\_revision 12-May-1995 #text\_change 09-Jul-2004  
 C.Accession: SS1424  
 R.Wohlmann, P.

submitted to the EMBL Data Library, November 1994  
 A.Description: The sequence of S. cerevisiae cosmid 9470.  
 A.Reference number: SS1414  
 A.Accession: SS1424  
 A.Molecule type: DNA  
 A.Residues: 1-201 <WOH>

Db 34 LP1KYGDLV 42

cytochrome P450 monooxygenase [imported] - Arabidopsis thaliana

C>Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 09-Jul-2004

C/Accession: T52173

R/Mizutani, M.; Ward, E.; Ohta, D.

Plant Mol. Biol. 37, 39-52, 1998

A/Title: Cytochrome P450 superfamily in Arabidopsis thaliana: isolation of cDNAs, differ

A/Reference number: 214382; PMID:98281573; PMID:9620263

A/Accession: T52173

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-502 <M1Z>

A/Cross-references: UNIPROT:P05788; EMBL:D78605; PDB:BA28537.1

A/Genetics:

C/Suprafamily: human cytochrome P450 CYP2D6; cytochrome P450 homology

C/Keywords: heme; iron; metalloprotein

F/301-466/Domain: cytochrome P450 homology <P45>

F/444/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 69.2%; Score 36; DB 2; Length 502;

Best Local Similarity 77.8%; Pred. No. 53;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LPIRYGPLY 10

DB 58 LPIRYGPLY 66

RESULT 13

peptide synthetase (EC 1.14.17.-) - imperfect fungus (Metarhizium anisopliae) (fragment)

C/Species: Metarhizium anisopliae

C/Date: 31-Dec-1996 #sequence\_revision 31-Dec-1996 #text\_change 09-Jul-2004

C/Accession: PC4198; PC4199

R/Balley, A.M.; Kershaw, M.J.; Hunt, B.A.; Paterson, I.C.; Charnley, A.K.; Reynolds, S.E

Gene 173, 195-197, 1996

A/Title: Cloning and sequence analysis of an intron-containing domain from a peptide syn

A/Reference number: PC4198; PMID:57082966; PMID:6864498

A/Accession: PC4198

A/Molecule type: DNA

A/Residues: 1-1078 <BAI>

A/Cross-references: UNIPROT:Q01135; EMBL:X89442; NID:G1015536; PDB:G1015537

A/Accession: PC4199

A/Molecule type: Protein

A/Residues: 58-67/136-146 <B12>

C/Genetics:

A/Genes: pes

A/Introns: 386/3; 762/2

C/Keywords: carrier protein; oxidoreductase

F/32-499/Domain: acetate-CoA ligase homology <ACLD>

F/526-594/Domain: acyl carrier protein homology <ACPP>

Query Match 69.2%; Score 36; DB 2; Length 1078;

Best Local Similarity 77.8%; Pred. No. 1,2e+02;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LPIRYGPLY 10

DB 729 LPIRYGPLY 737

RESULT 14

142524

A/ORE-M protein - vaccinia virus (strains WR and Copenhagen)

N/Alternate names: SALB protein

C/Species: vaccinia virus

A/Note: host Homo sapiens (man)

C/Date: 09-Nov-1990 #sequence\_revision 09-Nov-1990 #text\_change 09-Jul-2004

C/Accession: 142524; J01819

R/Johnson, G.P.

submitted to GenBank, June 1990

A/Reference number: A33172

A/Accession: 142524

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-76 <J0H>

A/Cross-references: UNIPROT:P20522

R/Smith, G.L.; Chan, Y.S.; Howard, S.T.

J. Gen. Virol. 72, 1349-1376, 1991

A/Title: Nucleotide sequence of 42kbp of vaccinia virus strain WR from near the right ir

A/Reference number: J01867; PMID:91259063; PMID:2045793

A/Accession: J01819

A/Molecule type: DNA

A/Residues: 1-76 <SMI>

A/Cross-references: DDBJ:D11079

Query Match 67.3%; Score 35; DB 2; Length 76;

Best Local Similarity 75.0%; Pred. No. 10;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LPIRYGPLY 9

DB 17 LPIRYGPLY 24

RESULT 15

F69166

hypothetical protein MTH506 - Methanobacterium thermoautotrophicum (strain Delta H)

C/Species: Methanobacterium thermoautotrophicum

C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004

C/Accession: F69166

R/Smith, D.R.; Doucette-Stamm, L.A.; Delonghery, C.; Lee, H.; Dubois, J.; Aldredge, T.;

Olun, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N

Xi, S.; Chun, G.M.; Daniels, C.D.; Mao, J.; Rice, S.; Noelling, U.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct

A/Reference number: A69000; PMID:98037514; PMID:9371463

A/Accession: F69166

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-216 <MTH>

A/Cross-references: UNIPROT:Q26606; GB:AE000834; GB:AE000666; NID:G2621574; PDB:AAE501

A/Experimental source: strain Delta H

C/Genetics:

A/Genes: MTH506

A/Start codon: GTG

Query Match 67.3%; Score 35; DB 2; Length 216;

Best Local Similarity 71.4%; Pred. No. 33;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIRYGP 8

DB 143 VPIRYGP 149

Search completed: November 30, 2004, 07:41:24

Job time : 21 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 ; Search time 76.4103 seconds

(without alignments)  
75.301 Million cell updates/sec

Title: US-10-008-377a-6

Sequence: 1 ALPIEXGPLV 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database: UniProt\_02:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	803	2 Q8TDJ5	Q8TDJ5 homo sapien
2	52	100.0	1620	1 ALK_HUMAN	Q8UM73 homo sapien
3	49	100.0	1621	1 ALK_MOUSE	P97793 mus musculu
4	44	84.6	130	2 Q7MM63	Q7MM63 vibrio vuln
5	44	84.6	130	2 Q8D8A5	Q8D8A5 vibrio vuln
6	42	80.8	130	2 Q8D8A2	Q8D8A2 vibrio para
7	41	78.8	116	2 Q7C771	Q7C771 agrobacteri
8	41	78.8	116	2 Q8UE11	Q8UE11 agrobacteri
9	41	78.8	339	2 Q8XXA2	Q8XXA2 raltecona s
10	40	76.9	529	2 Q7PV84	Q7PV84 anopheles g
11	40	76.9	1721	2 Q9UNR3	Q9UNR3 pacificastac
12	40	76.9	3277	2 Q9UNR4	Q9UNR4 gloobacter
13	39	75.0	144	2 Q8FN67	Q8FN67 corynebacte
14	38	73.1	161	2 Q8SD08	Q8SD08 pseudomonas
15	38	73.1	163	2 Q8KRD9	Q8KRD9 vibrio chol
16	38	73.1	401	2 Q8HXQ8	Q8HXQ8 pseudomonas
17	37	71.2	61	2 Q8ILR2	Q8ILR2 plasmodium
18	37	71.2	201	1 YL79_YEAST	Q8G255 saccharomyc
19	37	71.2	239	2 Q89M29	Q89M29 bradyrhizob
20	37	71.2	306	2 Q53929	Q53929 streptomyce
21	37	71.2	306	2 Q8CUC3	Q8CUC3 streptomyce
22	37	71.2	673	2 Q7OEW1	Q7OEW1 glycine max
23	37	71.2	673	2 Q8E4788	Q8E4788 glycine m
24	37	71.2	1083	1 EMBR_MYCLE	Q8C8A9 mycobacteri
25	37	71.2	7525	2 Q8KIE0	Q8KIE0 streptomyce
26	36	69.2	167	2 Q8GCA1	Q8GCA1 citrobacter
27	36	69.2	203	2 Q03201	Q03201 saccharomyc
28	36	69.2	203	2 A6555974	A6555974 saccharom
29	36	69.2	233	2 Q915C9	Q915C9 pseudomonas
30	36	69.2	298	2 Q7X295	Q7X295 streptomyce
31	36	69.2	322	2 Q7Q062	Q7Q062 anopheles g

## ALIGNMENTS

## RESULT 1

Q8TDJ5 PRELIMINARY; PRT; 803 AA.

AC Q8TDJ5

DT 01-UN-2002 (TREMBLrel. 21, Created)

DT 01-UN-2002 (TREMBLrel. 21, Last sequence update)

DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)

DE TRK-fused gene/anaplastic large cell lymphoma kinase extra long form.

GN Name=TRF/ALK fusion;

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21940189; PubMed=11943732;

RA Hernandez L., Bea S., Bellosillo B., Pinol M., Falini B., Ott G.,

RA Pulford K., Rosenwald A., Morris S.W., Fernandez A., Santos E.,

RA Campo E.;

RT "Diversity of genomic breakpoints in TRF-ALK translocations in

RT anaplastic large cell lymphomas: identification of a new TRF-ALK(XL)

RT chimeric gene with transforming activity.";

RT Am. J. Pathol. 160:1487-1494(2002).

DR EMBL; AF390893; AAM17922.1; -

DR HSSP; Q62838; 1LUF.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.

DR GO; GO:0006740; F:transferase activity; IEA.

DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.

DR InterPro; IPR011009; Kinase-like.

DR InterPro; IPR000270; OPR\_PBI.

DR InterPro; IPR000719; Prot\_Kinase.

DR InterPro; IPR002011; ReceptTyknsil.

DR InterPro; IPR001245; Tyr\_Kinase.

DR InterPro; IPR008266; Tyr\_Kinase\_AS.

DR Pfam; PF00564; PBI.1.

DR Pfam; PF00069; Ekinase.1.

DR PRINTS; PR00109; TYRKINASE.

DR ProDom; PD000001; Prot\_Kinase; 1.

DR SMART; SM00666; PBI.1.

DR SMART; SM00219; TykC.1.

DR PROSITE; PS00107; PROTEIN KINASE ATP; UNKNOWN\_1.

DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.

DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.

DR PROSITE; PS00239; RECEPTOR TYR\_KIN\_IL; UNKNOWN\_1.

KW ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.

SEQUENCE 803 AA; 8908260495B695F CRC64;

Query Match 100.0%; Score 52; DB 2; Length 803;

Best Local Similarity 100.0%; Pred. No. 0.42;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIEXGPLV 10

DB 579 ALPIRYGPLY 568

RESULT 2  
ID ALK\_HUMAN STANDARD; PRT; 1620 AA.  
AC Q9UM73; Q9Y4K6;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE ALK tyrosine kinase receptor precursor (EC 2.7.1.112) (Amplastic  
lymphoma kinase) (CD246 antigen).  
CN Name:ALK;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID:9606;  
[1]  
NP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, TISSUE SPECIFICITY, AND  
RP GLYCOSYLATION.  
RX MEDLINE=97316773; PubMed=9174053; DOI=10.1038/57.000.1201062;  
RA Morris S.W., Naeve C.W., Mathew P., James P.L., Kirstein M.N., Cui X.,  
Witte D.P.;  
RT "ALK, the chromosome 2 gene locus altered by the t(2;5) in non-  
Hodgkin's lymphoma, encodes a novel neutral receptor tyrosine kinase  
that is highly related to leukocyte tyrosine kinase (LTK)."  
RL Oncogene 14:2175-2188(1997).  
[2]  
RN ERRATUM.  
RA Morris S.W., Naeve C.W., Mathew P., James P.L., Kirstein M.N., Cui X.,  
Witte D.P.;  
RL Oncogene 15:2883-2883(1997).  
[3]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=97178863; PubMed=9053841;  
RA Iwahara T., Fujimoto J., Wen D., Cupples R., Bucay N., Arakawa T.,  
Mori S., Ratakin B., Yamamoto T.;  
RT "Molecular characterization of ALK, a receptor tyrosine kinase  
expressed specifically in the nervous system."  
RL Oncogene 14:439-449(1997).  
[4]  
RN PARTIAL SEQUENCE FROM N.A., AND CHROMOSOMAL TRANSLOCATION.  
RX MEDLINE=94167588; PubMed=8122112;  
RA Morris S.W., Kirstein M.N., Valentine M.B., Dittmer K.G.,  
Shapiro D.N., Saltman D.L., Look A.T.;  
RT "Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in  
non-Hodgkin's lymphoma."  
RL Science 263:1261-1264(1994).  
CC -1- FUNCTION: Orphan receptor with a tyrosine-protein kinase activity.  
CC Appears to play an important role in the normal development and  
CC function of the nervous system.  
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- TISSUE SPECIFICITY: Expressed in brain and CNS. Also expressed in  
CC the small intestine and testis, but not in normal lymphoid cells.  
CC -1- PTM: N-glycosylated.  
CC -1- DISEASE: A form of non-Hodgkin's lymphoma is characterized by a  
CC chromosomal translocation t(2;5)(p23;q35) that involves NPM1 and  
CC ALK.  
CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin  
CC receptor subfamily.  
CC -1- SIMILARITY: Contains 1 LDL-receptor class A domain.  
CC -1- SIMILARITY: Contains 2 MAM domains.  
CC -1- DATABASE: NMRAtlas Genet. Cyrogenet. OncoJ. Haematol.;  
CC WWW="http://www.inobio.gen.fr/services/chromancer/Genes/ALK.html".  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL, U62540; AAB7619.1; -  
DR EMBL, U66559; AAC51104.1; -  
DR HSSP, Q62838; ILUF.  
DR Genew; HGNC:427; ALK.  
DR MIM; 105590; -  
DR GO; GO:0016021; C: integral to membrane; NAS.  
DR GO; GO:0007399; P: neurogenesis; NAS.  
DR InterPro; IPR011009; Kinase like.  
DR InterPro; IPR002172; LPL receptor\_A.  
DR InterPro; IPR000998; MAM\_1.  
DR InterPro; IPR000719; Prot. kinase.  
DR InterPro; IPR002011; ReceptTyknsf1.  
DR InterPro; IPR001245; Tyr. kinase.  
DR InterPro; IPR008266; Tyr. kinase\_AS.  
DR Pfam; PF00629; MAM; 1.  
DR Pfam; PF00069; Pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000601; Prot. kinase; 1.  
DR SMART; SM00192; LPLA; 1.  
DR SMART; SM01209; TYKCA; 1.  
DR PROSITE; PS01209; LPLA\_1; FALSE NEG.  
DR PROSITE; PS00068; LPLA\_2; FALSE NEG.  
DR PROSITE; PS00740; MAM\_1; FALSE NEG.  
DR PROSITE; PS00060; MAM\_2; 2.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00339; RECEPTOR\_TYR\_KIN\_IL1; 1.  
DR ATP-binding; Chromosomal translocation; Glycoprotein; Phosphorylation;  
DR Proto-oncogene; Receptor; Repeat; Signal; Transferase; Transmembrane;  
KM Tyrosine-protein kinase.  
KW SIGNAL  
FT 1 18  
FT CHAIN 19 1620  
FT DOMAIN 19 1038  
FT TRANSMEM 1039 1059  
FT DOMAIN 1060 1620  
FT DOMAIN 264 427  
FT DOMAIN 437 473  
FT DOMAIN 478 636  
FT DOMAIN 1116 1392  
FT DOMAIN 816 940  
FT NP BIND 1122 1130  
FT BINDING 1150 1150  
FT ACT SITE 1249 1249  
FT MOD\_RES 1282 1282  
FT CARBOHYD 169 169  
FT CARBOHYD 244 244  
FT CARBOHYD 285 285  
FT CARBOHYD 324 324  
FT CARBOHYD 411 411  
FT CARBOHYD 424 424  
FT CARBOHYD 445 445  
FT CARBOHYD 563 563  
FT CARBOHYD 571 571  
FT CARBOHYD 627 627  
FT CARBOHYD 709 709  
FT CARBOHYD 808 808  
FT CARBOHYD 863 863  
FT CARBOHYD 864 864  
FT CARBOHYD 886 886  
FT CARBOHYD 986 986  
FT CONFLICT 36 36  
FT CONFLICT 1491 1491  
FT CONFLICT 1529 1529  
SQ SEQUENCE 1620 AA; 176417 MW; A62604B2429E15LE CRC64;  
Query Match 100.0%; Score 52; DB 1; Length 1620;  
Best Local Similarity 100.0%; Pred. No. 0.87;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 ALPTEVGPV 10  
DB 1396 ALPTEVGPV 1405

RESULT 3  
ALK\_MOUSE STANDARD; PRT; 1621 AA.

AC P97793;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE ALK tyrosine kinase receptor precursor (EC 2.7.1.112) (Anaplastic  
DE Lymphoma kinase).  
GN Name=Alk;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain, and Testis;  
RX MEDLINE=9717863; PubMed=9053841;  
RA Iwahara T., Fujimoto J., Wen D., Cupples R., Bucay N., Arakawa T.,  
RA Mori S., Ratkai B., Yamamoto T.;  
RT "Molecular characterization of Alk, a receptor tyrosine kinase  
RT expressed specifically in the nervous system."  
RL Oncogene 14:439-449(1997).  
CC -1- FUNCTION: Orphan receptor with a tyrosine-protein kinase activity.  
CC Appears to play an important role in the normal development and  
CC function of the nervous system.  
CC Cytosolic ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Probable).  
CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin  
CC receptor subfamily.  
CC -1- SIMILARITY: Contains 1 LDL-receptor class A domain.  
CC -1- SIMILARITY: Contains 2 MAM domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D83002; BA01673.1; -;  
DR PIR; T30200; T30200.  
DR HSSP; P08063; L1QH.  
DR WGD; MG1403505; ALK.  
DR GO; GO:0005515; F:Protein binding; IPT.  
DR InterPro; IPR011008; Kinase\_1ike.  
DR InterPro; IPR002172; LDL\_receptor\_A.  
DR InterPro; IPR000998; MAM.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR002011; ReceptTyrosinI.  
DR InterPro; IPR001245; Tyr\_kinase.  
DR InterPro; IPR008266; Tyr\_kinase\_AS.  
DR Pfam; PF00629; MAM; 1.  
DR Pfam; PF00069; PKinase\_1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Proc\_kinase; 1.  
DR SMART; SM00192; Ldla; 1.  
DR SMART; SM00219; TYRK; 1.  
DR PROSITE; PS01209; LDLRA\_1; FALSE\_NEG.  
DR PROSITE; PS00068; LDLRA\_2; FALSE\_NEG.  
DR PROSITE; PS00740; MAM\_1; FALSE\_NEG.  
DR PROSITE; PS00060; MAM\_2; 2.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS0011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.

DR PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
KW ATP-binding; Glycoprotein; Phosphorylation; Receptor; Repeat; Signal;  
KW Transferase; Transmembrane; Tyrosine-protein kinase.  
FT SIGNAL 1  
FT CHAIN 18  
FT DOMAIN 19 1621  
FT TRANSMEM 19 1042  
FT TRANSMEM 1043 1063  
FT DOMAIN 1064 1621  
FT DOMAIN 268 431  
FT DOMAIN 441 477  
FT DOMAIN 482 640  
FT DOMAIN 1120 1396  
FT DOMAIN 944  
FT NP\_BIND 1126 1134  
FT BINDING 1154 1154  
FT ACT\_SITE 1253 1253  
FT MOD\_RES 1286 1286  
FT CARBOHYD 174 174  
FT CARBOHYD 248 248  
FT CARBOHYD 289 289  
FT CARBOHYD 328 328  
FT CARBOHYD 415 415  
FT CARBOHYD 428 428  
FT CARBOHYD 449 449  
FT CARBOHYD 567 567  
FT CARBOHYD 575 575  
FT CARBOHYD 631 631  
FT CARBOHYD 673 673  
FT CARBOHYD 713 713  
FT CARBOHYD 812 812  
FT CARBOHYD 868 868  
FT CARBOHYD 890 890  
FT CARBOHYD 990 990  
SQ SEQUENCE 1621 AA; 174919 MW; 16E25B21AADE22 CRC64;

Query Match 94.2%; Score 49; DB 1; Length 1621;  
Best Local Similarity 90.0%; Pred. No. 3.4;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

CY 1 ALPTEVGPV 10  
DB 1400 ALPTEVGPV 1409

RESULT 4  
ID Q7MM63 PRELIMINARY; PRT; 130 AA.  
AC Q7MM63;  
DT 01-MAR-2004 (TRENBERG; 26, Created)  
DT 01-MAR-2004 (TRENBERG; 26, Last sequence update)  
DT 01-MAR-2004 (TRENBERG; 26, Last annotation update)  
DE Hypothetical protein V1210.  
GN Name=V1210;  
OS Vibrio vulnificus (strain V1016).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrrio.  
OX NCBI\_TaxID=196600;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed14656965;  
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,  
RA Liao T.-L., Liu Y.-M., Chen H.-Y., Shen A.B.-T., Li J.-C., Su T.-L.,  
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;  
RT "Comparative genome analysis of Vibrio vulnificus, a marine  
RT pathogen."  
RL Genome Res. 13:2577-2587(2003).  
DR EMBL; AP005334; BAC93974.1; -;  
KM Hypothetical protein.  
SQ SEQUENCE 130 AA; 14171 MW; 6C1BC5D7EC734063 CRC64;

Query Match 84.6%; Score 44; DB 2; Length 130;  
Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIEXGPL 9  
DB 43 LPIEXGPL 50

## RESULT 5

Q8D8A5 PRELIMINARY; PRT; 130 AA.  
AC Q8D8A5;  
DT 01-MAR-2003 (TREMBlrel. 23, Created)  
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE Hypothetical protein.  
GN OrderedLocustNames=V13076;  
OS *Vibrio vulnificus*.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; *Vibrio*.  
OX NCBI\_TaxID=672;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CMP6;  
RA Rhee U.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,  
RA Choy H.E.,  
RT "Complete genome sequence of *Vibrio vulnificus* CMP6.";  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL, AE016807; AAO11400.1;  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 130 AA; 14203 MW; 6C1BC5DDEC9F1C8 CRC64;

Query Match 84.6%; Score 44; DB 2; Length 130;  
Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIEXGPL 9  
DB 43 LPIEXGPL 50

## RESULT 6

Q87NA2 PRELIMINARY; PRT; 130 AA.  
AC Q87NA2;  
DT 01-JUN-2003 (TREMBlrel. 24, Created)  
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE Hypothetical protein VPI973.  
GN OrderedLocustNames=VPI973;  
OS *Vibrio parahaemolyticus*.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; *Vibrio*.  
OX NCBI\_TaxID=670;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RMD 2210633 / Serotype O3:K6;  
RX MEDLINE=22508454; PubMed=12620739;  
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,  
RA Iijima Y., Nishina M., Nakano M., Yamashita A., Kubota Y., Kimura S.,  
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.,  
RT "Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism  
distinct from that of *V. cholerae*.";  
RL Lancet 361:743-749 (2003).  
DR EMBL, AP005079; BAC60236.1;  
DR InterPro: IPR01047; Quin a1c DH like.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 130 AA; 14412 MW; 7E3F1A9DE506AC CRC64;

Query Match 80.8%; Score 42; DB 2; Length 130;  
Best Local Similarity 87.5%; Pred. No. 5.7;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIEXGPL 9

DB 42 LPIEXGPL 49

Q7CY71 PRELIMINARY; PRT; 116 AA.

AC Q7CY71;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE AGR C 3564D.  
GN OrderedLocustNames=AGR C 3564;  
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
OX NCBI\_TaxID=176299;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Cereon;  
RX MEDLINE=21608551; PubMed=11743194;  
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
RA Curcio B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,  
RA Houniel K., Gordon J., Vaudin M., Iatchouk O., Epp A., Liu F.,  
RA Wollam C., Allinger M., Doughty D., Scott C., Jappes C., Markelz B.,  
RA Flanagan C., Crowell C., Gerson D., Lomo C., Seer C., Strub G.,  
RA Cielo C., Slater S.,  
RT "Genome sequence of the plant pathogen and biotechnology agent  
Agrobacterium tumefaciens C58.";  
RL Science 294:2323-2328 (2001).  
DR EMBL, AE008113; AAK8713.1;  
DR InterPro: IPR004350; Gly\_bleo\_diox.  
DR Pfam PF00903; Glyoxalase\_1.  
SQ SEQUENCE 116 AA; 12585 MW; 9FA61753836C7655 CRC64;

Query Match 78.8%; Score 41; DB 2; Length 116;  
Best Local Similarity 87.5%; Pred. No. 7.9;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIEXGPL 9  
DB 84 LPIEXGPL 91

## RESULT 8

Q8UE11 PRELIMINARY; PRT; 116 AA.  
AC Q8UE11;  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Hypothetical protein Adu1953.  
GN OrderedLocustNames=Adu1953;  
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
OX NCBI\_TaxID=176299;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Dupont;  
RX MEDLINE=21608550; PubMed=11743193;  
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.F., Woo L.,  
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,  
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D. Sr.,  
RA Chapman P., Clendenning J., Deatherage G., Gilliet W., Grant C.,  
RA Kutyavint T., Levy R., Li M.-J., McClelland E., Palmeri A.,  
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,  
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,  
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
RA Chunley F., Tingey S.V., Tomo J.-F., Gordon M.P., Olson M.V.,  
RA Nester E.W.,  
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens  
C58.";

```

RL Science 294:2317-2323(2001).
DR EMBL; A2009148; AAL42949.1; -.
DR PIR; AG2816; AG2816.
DR PIR; H97594; H97594.
DR InterPro; IPR004360; Gly_bleo_dtox.
DR Pfam; PF00903; Glyoxalase1.
KW Complete proteome, Hypothetical protein.
SQ SEQUENCE 116 AA; 12585 MW; 9FA6175386C7655 CRC64;

Query Match
Best Local Similarity 78.8%; Score 41; DB 2; Length 116;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPPIYGPL 9
DB 84 LPPIYGPLV 91

RESULT 9
Q8XXA2 PRELIMINARY; PRT; 339 AA.
AC Q8XXA2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein RSC2214
GN Name=RS01385; OrderedLocNames=RSC2214;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteriia; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM1000;
RX Medline=21681879; Pubmed=11823852;
RA Salanoubat M., Genin S., Arriaguenave F., Gouzy J., Mangenot S.,
RA Riatac M., Billault A., Brocher P., Camus J.C., Catolico L.,
RA Chandler M., Choise N., Claude-Renard C., Cumac S., Demange N.,
RA Gaupin C., Lavie W., Moisan A., Robert C., Saurin W., Schlex T.,
RA Sigulier P., Thebaud P., Whalen M., Winkler P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646068; CAD15921.1; -.
DR InterPro; IPR001279; Blackmase-1like.
DR Pfam; PF00753; Lactamase_B; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 339 AA; 36167 MW; 85684DE2A5DCDD26 CRC64;

Query Match
Best Local Similarity 78.8%; Score 41; DB 2; Length 339;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPPIYGPL 9
DB 326 LPPIYGPLV 333

RESULT 10
Q7PVB4 PRELIMINARY; PRT; 529 AA.
AC Q7PVB4;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP000001653 (Fragment).
GN Name=ENSANG000001653;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anophelinae.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008986; EAA00286.2; -.
DR InterPro; IPR002645; STAS.
DR Pfam; PF01740; STAS; 1.
DR Pfam; PF00916; Sulfate_transp; 1.
DR PROSITE; PS50801; STAS; 1.
FT NON TER
SQ SEQUENCE 529 AA; 57283 MW; 9A037334418341ED CRC64;

Query Match
Best Local Similarity 76.9%; Score 40; DB 2; Length 529;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALPIYGPLV 10
DB 382 ALPIYGPLV 391

RESULT 11
Q9UAR3 PRELIMINARY; PRT; 1721 AA.
AC Q9UAR3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Clotting protein precursor.
OS Pacifastacus leniusculus (Signal crayfish).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Decapoda; Decapoda; Pleocyemata; Astacidea;
OC Astacidea; Astacidae; Pacifastacus.
OX NCBI_TaxID=6720;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99162541; Pubmed=10051578;
RA Hall M., Wang R., van Amerongen R., Sottrup-Jensen L., Soderhall K.;
RT "The crayfish plasma clotting protein: a vitellogenin-related protein
RT responsible for clot formation in crustacean blood.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:1965-1970(1999).
DR EMBL; AF102268; AAD16454.1; -.
DR PIR; S30348; S30348.
DR GO; GO:0005319; F:lipid transporter activity; IEA.
DR GO; GO:0006869; F:lipid transport; IEA.
DR InterPro; IPR001747; Lipid_transport_N.
DR InterPro; IPR01030; LV_superhelical.
DR Pfam; PF01347; Vitellogenin_N; 1.
DR Pfam; PF00094; WMD; 1.
DR SMART; SMO0638; LPD_N; 1.
DR SMART; SMO0216; WMD; 1.
DR Signal.
DR Signal.
FT CHAIN
SQ SEQUENCE 1721 AA; 194665 MW; 5BD8CC9352C50CE7 CRC64;

Query Match
Best Local Similarity 76.9%; Score 40; DB 2; Length 1721;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LPPIYGPLV 10
DB 84 LPPIYGPLV 92

RESULT 12
Q7NNI4 PRELIMINARY; PRT; 3277 AA.
AC Q7NNI4;

```

DT 01-MAR-2004 (TREMBlrel. 26, Created)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE G110427 protein.  
 GN OrderedLocustNames=g110427;  
 OS Gloeobacter violaceus.  
 OC Bacteria; Cyanobacteria; Chroococcales; Gloeobacter.  
 NC NCB1\_TaxID=33072;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PCC 7421;  
 RX MEDLINE=22977040; PubMed=14621292;  
 RA Nakamura Y., Kaneko T., Sato S., Miumuro M., Miyashita H., Tsuchiya T.,  
 RA Koshino S., Matsumoto M., Matsuno A., Kishida Y., Kiyokawa C.,  
 RA Kohara M., Matsuno M., Matsuno A., Nakazaki N., Shimo S.,  
 RA Takemuchi C., Yamada M., Tabara S.;  
 RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a  
 RT cyanobacterium that lacks thylakoids."  
 RL DNA Res. 10:137-145(2003).  
 DR EMBL; AP06569; BAC8368.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 3277 AA; 337749 MW; BAF9CF2EDBBA5CD8 CRC64;

Query Match 75.9%; Score 40; DB 2; Length 3277;  
 Best Local Similarity 75.0%; Pred. No. 4e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 PIRYGPV 10  
 DB 2891 PIRYGPV 2898

RESULT 13  
 Q8RN67 PRELIMINARY; PRT; 144 AA.  
 AC Q8RN67;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocustNames=CE2280;  
 OS Corynebacterium efficiens.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.  
 NC NCB1\_TaxID=152794;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=YS-314;  
 RX MEDLINE=22723752; PubMed=12840036;  
 RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,  
 RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,  
 RA Gotohori T.;  
 RT "Comparative complete genome sequence analysis of the amino acid  
 RT replacements responsible for the thermostability of Corynebacterium  
 RT efficiens."  
 RL Genome Res. 13:1572-1579(2003).  
 DR EMBL; AP005221; BAC19090.1; -.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 144 AA; 16054 MW; 34162E296CAF4182 CRC64;

Query Match 75.0%; Score 39; DB 2; Length 144;  
 Best Local Similarity 85.7%; Pred. No. 24;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 PIRYGPV 9  
 DB 10 PIRYGPV 16

RESULT 14  
 Q8SD08 PRELIMINARY; PRT; 161 AA.  
 ID Q8SD08  
 AC Q8SD08;

DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE PHIKZ154.  
 GN Name=CRF154;  
 OS Pseudomonas phage phikz.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.  
 NC NCB1\_TaxID=169683;  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21914557; PubMed=11916376;  
 RA Mesyanzhinov V.V., Robben J., Grymoprez B., Kostychenko V.A.,  
 RA Boukaltsava M.V., Syktilinda N.N., Kiylov V.N., Volckart G.;  
 RT "The genome of bacteriophage phikz of Pseudomonas aeruginosa."  
 RL J. Mol. Biol. 317:1-19(2002).  
 DR EMBL; AF39011; ALB8305.1; -.  
 SQ SEQUENCE 161 AA; 18806 MW; 55520C3654704C89 CRC64;

Query Match 73.1%; Score 38; DB 2; Length 161;  
 Best Local Similarity 77.8%; Pred. No. 43;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 LPIRYGPV 10  
 DB 94 LPIRYGPV 102

RESULT 15  
 Q9KRD9 PRELIMINARY; PRT; 163 AA.  
 AC Q9KRD9;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Hypothetical protein VC1703.  
 GN OrderedLocustNames=VC1703;  
 OS Vibrio cholerae.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrio.  
 NC NCB1\_TaxID=566;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=El Tor N16961 / Serotype O1;  
 RX MEDLINE=20406833; PubMed=10952301; DOI=10.1038/35020000;  
 RA Heidelberg J.F., Eissen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,  
 RA Dodson R.U., Hail D.H., Hickey E.K., Peterson J.D., Nayam L.A.,  
 RA Gill S.R., Nelson K.E., Reed J.D., Retchin H., Richardson D.L.,  
 RA Ermolaeva M.D., Vamathevan J.U., Bass S., Qin H., Dragoi I.,  
 RA Sellers P., McDonald L.A., Utterback T.R., Fleischmann R.D.,  
 RA Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,  
 RA Mekalanos J.J., Venter J.C., Fraser C.M.;  
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
 RT cholerae."  
 RL Nature 406:477-483(2000).  
 DR EMBL; AE004246; AAF94853.1; -.  
 DR F1R; D82167; D82167.  
 DR TIGR; VC1703; -.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 163 AA; 17925 MW; FD33C0A29294B72 CRC64;

Query Match 73.1%; Score 38; DB 2; Length 163;  
 Best Local Similarity 75.0%; Pred. No. 43;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIRYGPV 9  
 DB 76 LPIRYGPV 83

Search completed: November 30, 2004, 07:17:30  
 Job time : 79.4103 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 ; Search time 77.1795 Seconds  
(without alignments)  
46,480 Million cell updates/sec

Title: US-10-008-377a-6  
Perfect score: 52  
Sequence: 1 ALPHEXGLV 10

Scoring table: BIOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 2002273 seqs, 35872929 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23sep04:\*

- 1: geneseqp1908:\*
- 2: geneseqp1908:\*
- 3: geneseqp2000:\*
- 4: geneseqp2001:\*
- 5: geneseqp2002:\*
- 6: geneseqp2003:\*
- 7: geneseqp2003:\*
- 8: geneseqp2004:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	10	AAO22988	Aao22988 Human p45
2	52	100.0	525	Aar78521	Aar78521 NPM/ALK f
3	52	100.0	525	ABU04368	ABU04368 Human exp
4	52	100.0	572	AAR78520	AAR78520 Partial A
5	52	100.0	675	ABG20950	ABG20950 Novel hum
6	52	100.0	680	AAR73858	Aar73858 Human NPM
7	52	100.0	680	ABG95067	ABG95067 Human tra
8	52	100.0	680	ABU04347	ABU04347 Human exp
9	52	100.0	680	ABU04362	ABU04362 Human exp
10	52	100.0	680	ABU04358	ABU04358 Human exp
11	52	100.0	680	ABG74472	ABG74472 NPM/ALK f
12	52	100.0	1620	AAI73857	AAI73857 Human ALK
13	52	100.0	1620	AAI018498	AAI018498 Human ins
14	52	100.0	1620	ABG74471	ABG74471 Human ALK
15	52	100.0	207	ADP06379	ADP06379 Bacterial
16	38	73.1	409	ABO69358	ABO69358 Pseudomon
17	37	71.2	58	AAR99175	AAR99175 Genetical
18	37	71.2	60	ADP71971	ADP71971 Kalikre
19	37	71.2	206	ABU04832	ABU04832 LDI recep
20	37	71.2	1083	ABU05679	ABU05679 M. tuberc
21	36	69.2	62	AAU61865	AAU61865 Proionib
22	36	69.2	62	ABMS8384	ABMS8384 Proionib
23	36	69.2	95	ABP05147	ABP05147 Human ORF
24	36	69.2	203	ABRS3343	ABRS3343 Protein s
25	36	69.2	203	ADK63424	ADK63424 Disease t

25	36	69.2	243	7	ABO80522	ABO80522 Pseudomon
27	36	69.2	326	8	ADP99000	ADP99000 C. albica
28	36	69.2	492	4	ABG28689	ABG28689 Novel hum
29	36	69.2	653	4	AAR37425	Aar37425 IDUA, 3/2
30	36	69.2	728	6	ABU18155	ABU18155 Protein e
31	35	67.3	51	2	AAR92258	Aar92258 TFP1-2 fi
32	35	67.3	53	2	AAR94976	Aar94976 Generic f
33	35	67.3	57	3	AAI14174	AAI14174 Tissue fa
34	35	67.3	58	2	AAR78547	Aar78547 Human TFP
35	35	67.3	58	2	AAR78548	Aar78548 Human TFP
36	35	67.3	58	2	AAR81920	Aar81920 Human TFP
37	35	67.3	58	2	AAR99173	Aar99173 Human apr
38	35	67.3	58	7	ADP42004	ADP42004 Human TFP
39	35	67.3	58	7	ADP71957	ADP71957 Kalikre
40	35	67.3	88	6	ABR39435	ABR39435 Human GEN
41	35	67.3	123	5	ABR98433	ABR98433 N-intein
42	35	67.3	123	7	ABM02089	ABM02089 Synchocy
43	35	67.3	123	7	ABM02087	ABM02087 Synchocy
44	35	67.3	123	7	ADL01661	ADL01661 Synchocy
45	35	67.3	123	7	ADL01639	ADL01639 Synchocy

## ALIGNMENTS

RESULT 1	AAO22988	standard; peptide; 10 AA.
ID	AAO22988	
AC	AAO22988	
DT	17-SEP-2003	(first entry)
DE	Human p45e-65 ALK-derived HLA-A*0201 restricted CTL epitope peptide.	
XX	HLA-A*0201 binding; anaplastic lymphoma kinase; ALK; CTL epitope; ALCL;	
KM	cytotoxic T-lymphocyte; receptor tyrosine; NPM/ALK fusion protein; human;	
KM	oncogenic; anaplastic large cell lymphoma; neuroblastoma; gene therapy;	
KM	t(2;5) translocation; neoplasia; cytostatic; immunotherapy; alkoma;	
XX	p45e-65.	
OS	Homo sapiens.	
XX		
PN	WO2003042243-A2.	
PD	22-MAY-2003.	
XX		
PF	14-NOV-2002; 2002WO-EP012764.	
XX		
PR	15-NOV-2001; 2001US-00008377.	
PA	(NAST-) 1ST NAZ STUDIO E CURA DEI TUMORI.	
PA	(SUJD-) ST JUDE CHILDREN'S RES HOSPITAL.	
XX		
PI	Gambacorti-Passerini C, Passoni L;	
XX		
DR	WPI; 2003-441791/41.	
XX		
PT	New HLA-A*0201-binding Anaplastic Lymphoma Kinase (ALK) peptide;	
PT	useful for preparing a composition for treating ALK-positive lymphoma,	
XX	neuroblastoma or ALK-expressing neoplasia.	
XX		
PS	Claim 1; Page 7; 33pp; English.	
XX		
CC	The invention relates to a novel HLA-A*0201-binding anaplastic lymphoma	
CC	kinase (ALK)-derived peptide, of 9-10 amino acids, which may act as a	
CC	cytotoxic T-lymphocyte (CTL)-epitope. ALK is a receptor tyrosine kinase	
CC	which forms part of the NPM/ALK fusion protein (oncogenic lymphoma kinase	
CC	fusion protein) in ALCL (anaplastic large cell lymphoma)-derived cell	
CC	lines. More than 50% of ALCL cases possess a t(2;5) chromosomal	
CC	translocation that leads to the expression of the NPM/ALK fusion protein	
CC	which forms a potent oncogene when constitutively activated. Translocated	
CC	ALK is a widely expressed tumour-associated antigen characteristic of ALK	

CC -positive lymphomas, neuroblastomas and ALK-expressing neoplasias. The  
 CC peptides of the invention demonstrate cytostatic activity and induce an  
 CC MHC (major histocompatibility complex) class I restricted cytotoxic  
 CC lymphocyte response against tumour cells expressing the NPM/ALK fusion  
 CC protein. Hence, the peptides may be utilised during the treatment, via  
 CC immunotherapy, of ALK-positive lymphomas such as alkomas, neuroblastomas  
 CC and ALK-expressing neoplasias. In addition, the peptides may be used  
 CC during gene therapy. The current sequence is that of the human p456-65  
 CC ALK-derived HLA-A\*0201 restricted CTL epitope peptide of the invention  
 XX

Sequence 10 AA;  
 SQ

Query Match 100.0%; Score 52; DB 6; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0064;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALPIEYGPLV 10  
 |||||  
 Db 1 ALPIEYGPLV 10

RESULT 2  
 AAR78521  
 ID AAR78521 standard; protein; 525 AA.  
 XX  
 AC AAR78521;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 15-NOV-1995 (first entry)  
 XX  
 DE NPM/ALK fusion protein.  
 XX  
 KW ALK protein; protein-tyrosine-kinase; fusion protein; NPM protein;  
 KW nuclear phosphoprotein; anaplastic large cell lymphoma; t(2;  
 XX 5) translocation; diagnosis; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO9515331-A1.  
 XX  
 PD 08-JUN-1995.  
 XX  
 PF 05-DEC-1994; 94WO-US013947.  
 XX  
 PR 03-DEC-1993; 93US-00160861.  
 XX  
 PA (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 XX  
 PI Morris SW, Look AT;  
 XX  
 DR WPI; 1995-215226/28.  
 DR N-PEDB; AAQ95542.  
 XX  
 PT Methode for detecting human t(2;5) lymphoma - for detection and diagnosis  
 PT of anaplastic large cell lymphoma(e).  
 XX  
 PS Claim 16; Page 42-44; 70pp; English.  
 XX  
 CC The translocation event that occurs in human t(2;5) lymphoma brings  
 CC sequences from the nuclear phosphoprotein (NPM) gene on chromosome 5q35  
 CC to those from a protein-tyrosine-kinase (ALK) gene on chromosome-2q23.  
 CC Nucleic acids encoding a partial ALK protein (AAR78520) and the ALK/NPM  
 CC fusion protein (AAR78521) were isolated (AAQ95541, AAQ95542).  
 CC Identification of the NPM/ALK fusion allows lymphoma diagnosis, or  
 CC therapy through the use of antisense RNA or ribozymes. (Updated on 25-MAR  
 CC -2003 to correct PN field.)  
 XX  
 XX Sequence 525 AA;  
 SQ

Query Match 100.0%; Score 52; DB 2; Length 525;  
 Best Local Similarity 100.0%; Pred. No. 0.49;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALPIEYGPLV 10  
 |||||  
 Db 456 ALPIEYGPLV 465

RESULT 3  
 ABU04368  
 ID ABU04368 standard; protein; 525 AA.  
 XX  
 AC ABU04368;  
 XX  
 DT 29-JAN-2003 (first entry)  
 DT  
 DE Human expressed protein tag (EPT) #1034.  
 XX  
 KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
 KW protease inhibitor; transporter; cytoskeletal protein;  
 KW receptor; transcription factor; cancer; MHC;  
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200278524-A2.  
 XX  
 PD 10-OCT-2002.  
 XX  
 PF 28-MAR-2002; 2002WO-US009671.  
 XX  
 PR 28-MAR-2001; 2001US-0279495P.  
 PR 21-MAY-2001; 2001US-0292544P.  
 PR 08-AUG-2001; 2001US-0310801P.  
 PR 01-OCT-2001; 2001US-0326370P.  
 PR 04-DEC-2001; 2001US-0336780P.  
 PR 20-FEB-2002; 2002US-0358985P.  
 XX  
 PA (ZYCO-) ZYCOs INC.  
 XX  
 PI Chicz RM, Tomlinson AJ, Urban RG;  
 XX  
 DR WPI; 2003-040607/03.  
 XX  
 PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 PT cytoskeletal proteins, receptors or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PT leukemia.  
 XX  
 XX Example 2; SEQ ID NO 1034; 134pp; English.  
 XX  
 PS The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide. The purified polypeptide, or the antibody that binds to this  
 CC polypeptide, is useful for treating cancer. The polypeptide is also  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptides and  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
 CC lymphoma or leukaemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: This sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 525 AA;  
 SQ

Query Match 100.0%; Score 52; DB 6; Length 525;  
 Best Local Similarity 100.0%; Pred. No. 0.49;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIEXGPLV 10  
 |||||  
 DB 456 ALPIEXGPLV 465

## RESULT 4

AAR78520  
 ID AAR78520 standard; protein; 572 AA.

AC AAR78520;

DT 25-MAR-2003 (revised)  
 DT 15-NOV-1995 (first entry)

DE Partial ALK protein.

KM ALK protein; protein-tyrosine-kinase; fusion protein; NPM protein;  
 KM nuclear phosphoprotein; anaplastic large cell lymphoma; t(2;  
 KM 5) translocation; diagnosis; gene therapy.

OS Homo sapiens.

PN WO9515331-A1.

PD 08-JUN-1995.

PF 05-DEC-1994; 94WO-US013947.

PR 03-DEC-1993; 93US-00160861.

PA (SUD-) ST JUDE CHILDREN'S RES HOSPITAL.

PI Morris SW, Look AT;

XX WPI; 1995-215226/28.

DR N-PSDB; AAQ95541.

PT Methods for detecting human t(2;5) lymphoma - for detection and diagnosis

XX of anaplastic large cell lymphoma(s).

XX Claim 21; Page 40-42; 70pp; English.

XX The translocation event that occurs in human t(2;5) lymphoma brings

CC sequences from the nuclear phosphoprotein (NPM) gene on chromosome 5q35

CC to those from a protein-tyrosine-kinase (ALK) gene on chromosome 2q23.

CC Nucleic acids encoding a partial ALK protein (AAR78520) and the ALK/NPM

CC fusion protein (AAR78521) were isolated (AAQ95541, AAQ95542).

CC Identification of the NPM/ALK fusion allows lymphoma diagnosis, or

CC therapy through the use of antisense RNA or ribozymes. (Updated on 25-MAR

CC -2003 to correct PN field.)

XX SQ Sequence 572 AA;

QY Query Match 100.0%; Score 52; DB 2; Length 572;

Best Local Similarity 100.0%; Pred. No. 0.54;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 503 ALPIEXGPLV 512

RESULT 5

ABG20950

ID ABG20950 standard; protein; 675 AA.

AC ABG20950;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #20941.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KM food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

OS WO200175067-A2.

XX 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US008631.

PF 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

PA Drmanac RT, Liu C, Tang YT;

PI WPI; 2001-639362/73.

DR N-PSDB; AAS85137.

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

PS Claim 20; SEQ ID NO 51309; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)

CC sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping;

CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (II) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is

CC useful for generating antibodies against it, detecting or quantitating a

CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activity. The

CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG20950-ABG30377 represent novel human diagnostic

CC amino acid sequences of the invention. Note: The sequence data for this

CC patent did not appear in the printed specification, but was obtained in

CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 675 AA;

QY Query Match 100.0%; Score 52; DB 4; Length 675;

Best Local Similarity 100.0%; Pred. No. 0.65;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 451 ALPIEXGPLV 460

RESULT 6

AAAB73858

ID AAAB73858 standard; protein; 680 AA.

AC AAAB73858;

DT 15-MAY-2001 (first entry)

DE Human NPM/ALK fusion protein.

XX Human; anaplastic lymphoma kinase; ALK; nucleophosmin; NPM;

XX chromosomal translocation; cancer; NPM/ALK; fusion.

OS Homo sapiens.  
 XX US6174674-B1.  
 XX 16-JAN-2001.  
 PD 19-JUN-1998; 98US-00100089.  
 XX 03-DEC-1993; 93US-00160861.  
 PR 12-OCT-1995; 95US-00542363.  
 XX (SUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 PA Morris SW, Look AT;  
 PI WPI; 2001-243208/25.  
 DR N-PSDB; AAF76868.  
 XX  
 XX Detection of chromosomal rearrangement or translocations present in t(2;  
 PT 5) for cancer diagnosis comprises using and identifying anaplastic  
 PT lymphoma kinase and nucleophosmin genes in nucleic acid hybridization  
 PT techniques.  
 XX  
 XX Claim 1; Fig 2A; 87bp; English.  
 PS  
 XX The present sequence is given in a specification relating to a method for  
 CC detecting a chromosomal rearrangement involving a breakpoint in the  
 CC anaplastic lymphoma kinase (ALK) or nucleophosmin (NPM) gene. The method  
 CC is based on the identification and sequence determination of a novel  
 CC gene, ALK, which is fused to the gene encoding NPM in translocations  
 CC present in t(2;5) lymphoma cells. The method is useful for detecting a  
 CC t(2;5) chromosomal rearrangement, or a chromosomal rearrangement  
 CC involving the ALK gene or NPM gene. It is particularly useful in  
 CC determining if particular cells or tissues express ALK or NPM/ALK coding  
 CC sequences. The method is also useful in diagnostic assays to determine,  
 CC for example, if a mammal has cancer or a genetic predisposition to cancer  
 XX  
 SQ Sequence 680 AA;

Query Match 100.0%; Score 52; DB 4; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 0.65;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIEYGPLV 10  
 |||||  
 DB 456 ALPIEYGPLV 465

RESULT 7  
 ID ABG95067 standard; protein; 680 AA.  
 AC ABG95067;

DT 04-DEC-2002 (first entry)

DE Human translocation (2; 5) (p23; q35) protein.

XX Chromosome aberration; oncogenic fusion protein; cancer;  
 XX Proliferative disease; cellular protein isoform; heat shock protein 90;  
 KW HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;  
 KW T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;  
 KW acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;  
 KW acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;  
 KW papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;  
 KW rhabdomyosarcoma; synovial sarcoma; viral infection.

OS Homo sapiens.

XX WO200269900-A2.

PD 12-SEP-2002.

PF 01-MAR-2002; 2002WO-US060518.  
 XX 01-MAR-2001; 2001US-0272751P.  
 XX (CONF-) CONFORMA THERAPEUTICS CORP.  
 PA Fritz LC, Burrows FJ;  
 PI WPI; 2002-698710/75.  
 DR N-PSDB; ABS73246.  
 XX  
 XX Treating genetically-defined disease associated with chromosomal  
 PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative  
 PT diseases, involves administering an inhibitor of heat shock protein 90.  
 XX  
 XX Disclosure; Page 200-202; 389pp; English.  
 PS  
 XX The invention describes a method of treating genetically-defined disease  
 CC associated with chromosomal aberrations yielding oncogenic fusion  
 CC proteins (I), treating cancerous cells containing (I) in a heterogeneous  
 CC cell population, treating proliferative diseases associated with mutant  
 CC protein or cellular protein isoforms (II) dependent on heat shock protein  
 CC (HSP)-90, or selectively treating cells expressing (II) involving  
 CC administering HSP90-inhibitor. The method is useful for treating  
 CC genetically-defined disease with chromosomal aberration yielding  
 CC oncogenic fusion protein, treating cancerous cells containing fusion  
 CC protein in heterogeneous cell population, treating proliferative disease  
 CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or  
 CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.  
 CC p53), or selectively treating cells expressing mutant protein or cellular  
 CC protein isoform in a patient heterozygous for (II). The method is useful  
 CC for treating a disease e.g. haematopoietic disorder such as T or B cell  
 CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, NHL and CMML,  
 CC or a disease characterised by a solid tumour such as papillary thyroid  
 CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and  
 CC synovial sarcoma. The method is also useful for treating viral  
 CC infections. This represents a protein encoded by the DNA sequence of a  
 CC chromosome aberration  
 XX  
 SQ Sequence 680 AA;

Query Match 100.0%; Score 52; DB 5; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 0.65;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIEYGPLV 10  
 |||||  
 DB 456 ALPIEYGPLV 465

RESULT 8  
 ID ABU04347 standard; protein; 680 AA.  
 AC ABU04347;

DT 29-JAN-2003 (first entry)

DE Human expressed protein tag (EPT) #1013.

XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
 KW protease; protease inhibitor; transporter; cytoskeletal protein;  
 KW receptor; transcription factor; cancer; MHC;  
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

OS Homo sapiens.

XX WO200278524-A2.

XX 10-OCT-2002.

PF 28-MAR-2002; 2002WO-US009671.

XX 28-MAR-2001; 2001US-0279495F.  
PR 21-MAY-2001; 2001US-0292444P.  
PR 08-AUG-2001; 2001US-0310801P.  
PR 01-OCT-2001; 2001US-0326370P.  
PR 04-DEC-2001; 2001US-0336780P.  
PR 20-FEB-2002; 2002US-0358985P.  
XX  
PA (ZYCO-) ZYCOS INC.  
XX  
PI Chicx RM, Tomlinson AJ, Urban RG,  
XX WPI; 2003-040607/03.  
XX  
DR WPI; 2003-040607/03.  
XX  
PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
PT cytoskeletal proteins, receptors or transcription factors), useful for  
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
PT leukemia.  
XX  
PS Example 2; SEQ ID NO 1013; 134pp; English.  
XX  
CC The invention describes a purified polypeptide, which comprises a  
CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
CC transporter, cytoskeletal protein, receptor or transcription factor. The  
CC polypeptide is useful as an immunogenic composition for eliciting in a  
CC mammal an immunogenic response directed against any of the purified  
CC polypeptide. The purified polypeptide, or the antibody that binds to this  
CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 680 AA;  
Query Match 100.0%; Score 52; DB 6; Length 680;  
Best Local Similarity 100.0%; Pred. No. 0.65;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ALPIEYGPLV 10  
DB 456 ALPIEYGPLV 465  
RESULT 9  
ABU04362  
ID ABU04362 standard; protein; 680 AA.  
XX  
AC ABU04362;  
XX  
DT 29-JAN-2003 (first entry)  
XX  
DE Human expressed protein tag (EPT) #1028.  
XX  
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200278524-A2.  
XX  
PD 10-OCT-2002.  
XX  
PF 28-MAR-2002; 2002WO-US009671.

XX 28-MAR-2001; 2001US-0279495F.  
PR 21-MAY-2001; 2001US-0292444P.  
PR 08-AUG-2001; 2001US-0310801P.  
PR 01-OCT-2001; 2001US-0326370P.  
PR 04-DEC-2001; 2001US-0336780P.  
PR 20-FEB-2002; 2002US-0358985P.  
XX  
PA (ZYCO-) ZYCOS INC.  
XX  
PI Chicx RM, Tomlinson AJ, Urban RG,  
XX WPI; 2003-040607/03.  
XX  
DR WPI; 2003-040607/03.  
XX  
PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
PT cytoskeletal proteins, receptors or transcription factors), useful for  
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
PT leukemia.  
XX  
PS Example 2; SEQ ID NO 1028; 134pp; English.  
XX  
CC The invention describes a purified polypeptide, which comprises a  
CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
CC transporter, cytoskeletal protein, receptor or transcription factor. The  
CC polypeptide is useful as an immunogenic composition for eliciting in a  
CC mammal an immunogenic response directed against any of the purified  
CC polypeptide. The purified polypeptide, or the antibody that binds to this  
CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 680 AA;  
Query Match 100.0%; Score 52; DB 6; Length 680;  
Best Local Similarity 100.0%; Pred. No. 0.65;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ALPIEYGPLV 10  
DB 456 ALPIEYGPLV 465  
RESULT 10  
ABU04358  
ID ABU04358 standard; protein; 680 AA.  
XX  
AC ABU04358;  
XX  
DT 29-JAN-2003 (first entry)  
XX  
DE Human expressed protein tag (EPT) #1024.  
XX  
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200278524-A2.  
XX  
PD 10-OCT-2002.  
XX  
PF 28-MAR-2002; 2002WO-US009671.

XX 28-MAR-2001; 2001US-0279495P.  
 PR 21-MAY-2001; 2001US-0292544P.  
 PR 08-AUG-2001; 2001US-0310801P.  
 PR 01-OCT-2001; 2001US-0326370P.  
 PR 04-DEC-2001; 2001US-0336780P.  
 PR 20-FEB-2002; 2002US-0358955P.

XX (ZYCO-) ZYCOs INC.

XX Chicz RM, Tomlinson AJ, Urban RG;

XX WPI; 2003-040607/03.

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 PT cytoskeletal proteins, receptor or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PT leukemia.

XX Example 2; SEQ ID NO 1024; 134bp; English.

XX The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide. The purified polypeptide or the antibody that binds to this  
 CC polypeptide, is useful for treating cancer. The polypeptide is also  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptides and  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
 CC lymphoma or leukemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: This sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 680 AA;

Query Match 100.0%; Score 52; DB 6; Length 680;

Best Local Similarity 100.0%; Pred. No. 0.65; Mismatches 0; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIEYGPLV 10

DB 456 ALPIEYGPLV 465

RESULT 11

ID ABG74472 standard; protein; 680 AA.

XX ABG74472;

XX 11-APR-2003 (first entry)

XX NPM/ALK fusion construct SEQ ID 4.

XX ALK; human; nucleophosmin; NPM; anaplastic lymphoma kinase; chromosome 5;

XX t(2;5) lymphoma cell; chromosomal rearrangement; chromosome 2;

XX nucleolar phosphoprotein; centromeric; telomeric.

XX Synthetic.

XX Key

XX Protein

XX Domain

XX Protein

XX

XX

XX

FT Domain 182..437

FT /label= ALK\_catalytic\_domain

XX US6451997-B1.

XX 17-SEP-2002.

XX 28-SEP-2000; 2000US-00670827.

XX 03-DEC-1993; 93US-00160861.

XX 12-OCT-1995; 95US-00542363.

XX 19-JUN-1998; 98US-00100089.

XX (SUJD-) ST JUDE CHILDREN'S RES HOSPITAL.

XX Morris SM, Look AT;

XX WPI; 2003-101739/09.

XX N-PSDB; ABQ77107.

XX Kit for use in method of detecting t(2;5) chromosomal rearrangements or

XX rearrangements involving nucleolar phosphoprotein (NPM) or anaplastic

XX lymphoma kinase (ALK) comprises probes that hybridizes with NPM and ALK

XX genes.

XX Example 1; Fig 2A; 87bp; English.

XX This invention describes a novel kit for use in a method of detecting t(2

XX /5) chromosomal rearrangements (CR) or CRs involving nucleolar

XX phosphoprotein (NPM) and anaplastic lymphoma kinase (ALK). The method

XX comprises labelling probes that hybridise to a sequence of 2442 or 6226

XX nucleotides or their complement, and to region of human chromosome 5/2

XX having wild-type NPM/ALK genes (the region is centromeric/telomeric to

XX nucleotide). This sequence represents a polypeptide sequence described in

XX the disclosure of the invention

XX Sequence 680 AA;

Query Match 100.0%; Score 52; DB 6; Length 680;

Best Local Similarity 100.0%; Pred. No. 0.65; Mismatches 0; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIEYGPLV 10

DB 456 ALPIEYGPLV 465

RESULT 12

ID AAB73857 standard; protein; 1620 AA.

XX AAB73857;

XX 15-MAY-2001 (first entry)

XX Human ALK protein.

XX Human; anaplastic lymphoma kinase; ALK; nucleophosmin; NPM;

XX chromosomal translocation; cancer.

XX Homo sapiens.

XX US6174674-B1.

XX 16-JAN-2001.

XX 19-JUN-1998; 98US-00100089.

XX 03-DEC-1993; 93US-00160861.

XX 12-OCT-1995; 95US-00542363.

XX (SUJD-) ST JUDE CHILDREN'S RES HOSPITAL.

PT Morris SW, Look AT;  
 XX WPI; 2001-243208/25.  
 DR N-PSDB; AAF76867.  
 XX  
 XX  
 PT Detection of chromosomal rearrangement or translocations present in t(2;  
 PT 5) for cancer diagnosis comprises using and identifying anaplastic  
 PT lymphoma kinase and nucleophosmin genes in nucleic acid hybridization  
 PT techniques.  
 XX  
 XX  
 XX Claim 10; Fig 3B; 87pp; English.  
 CC  
 CC The present sequence is given in a specification relating to a method for  
 CC detecting a chromosomal rearrangement involving a breakpoint in the  
 CC anaplastic lymphoma kinase (ALK) or nucleophosmin (NPM) gene. The method  
 CC is based on the identification and sequence determination of a novel  
 CC gene, ALK, which is fused to the gene encoding NPM in translocations  
 CC present in t(2;5) lymphoma cells. The method is useful for detecting a  
 CC t(2;5) chromosomal rearrangement, or a chromosomal rearrangement  
 CC involving the ALK gene or NPM gene. It is particularly useful in  
 CC determining if particular cells or tissues express ALK or NPM/ALK coding  
 CC sequences. The method is also useful in diagnostic assays to determine,  
 CC for example, if a mammal has cancer or a genetic predisposition to cancer  
 XX  
 XX Sequence 1620 AA;  
 SQ  
 Query Match 100.0%; Score 52; DB 4; Length 1620;  
 Best Local Similarity 100.0%; Pred. No. 1.7;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ALPIRYGPIV 10  
 DB 1396 ALPIRYGPIV 1405  
 RESULT 13  
 ID AAO18498  
 AC AAO18498 standard; protein; 1620 AA.  
 XX  
 XX AAO18498;  
 XX  
 DT 11-OCT-2002 (first entry)  
 XX  
 DS Human insulin receptor signaling modifier SEQ ID NO: 18.  
 XX  
 KM Human; insulin receptor signaling; insulin receptor signaling modifier;  
 KM IHW; diabetes; metabolic syndrome; antidiabetic.  
 XX  
 OS Homo sapiens.  
 XX  
 WO200255664-A2.  
 PN  
 PN 18-JUL-2002.  
 PD  
 PF 11-JAN-2002; 2002WO-US001048.  
 XX  
 XX  
 PR 12-JAN-2001; 2001US-0261226P.  
 PR 12-JAN-2001; 2001US-0261303P.  
 PR 12-JAN-2001; 2001US-0261304P.  
 PR 12-JAN-2001; 2001US-0261335P.  
 PR 12-JAN-2001; 2001US-0261361P.  
 PR 12-JAN-2001; 2001US-0261456P.  
 PR 12-JAN-2001; 2001US-0261457P.  
 PR 12-JAN-2001; 2001US-0261458P.  
 PR 12-JAN-2001; 2001US-0261459P.  
 PR 12-JAN-2001; 2001US-0261461P.  
 PR 12-JAN-2001; 2001US-0261518P.  
 PR 12-JAN-2001; 2001US-0261531P.  
 PR 12-JAN-2001; 2001US-0261532P.  
 PR 12-JAN-2001; 2001US-0261589P.  
 PR 12-JAN-2001; 2001US-0261590P.  
 PR 12-JAN-2001; 2001US-0261694P.  
 DR

PR	12-JAN-2001;	2001US-0261695P.
PR	12-JAN-2001;	2001US-0261697P.
XX	(EXEL-) EXELIXIS INC.	
XX	Seidel-Dugan C,	Ferguson KC, Kidd T;
XX	WPI; 2002-599664/64.	
DR	N-PDB;	AAI48617.
PT	Identifying an insulin receptor signaling modulator, useful as drug targets for treating diabetes or metabolic disorders, comprises contacting an assay system comprising insulin receptor signaling modifiers with a test agent.	
XX	Disclosure; Page 59-66; 232pp; English.	
CC	The present invention relates to a method of identifying a candidate insulin receptor (INR) signaling modulating agent, involving contacting an assay system comprising an insulin receptor signaling modifier (ISM) polypeptide or nucleic acid with a test agent, and detecting a test agent related activity of the assay system. The method is useful for identifying candidate INR signaling modulating agents. ISM genes may be used as drug targets for treatment of disorders related to INR signaling such as diabetes or metabolic syndrome. ISM nucleic acids and polypeptides are useful for identifying and testing agents that modulate INR function and for other applications related to the involvement of ISM in INR signaling, and for identifying subjects having a predisposition to such diseases associated with INR signaling. The present sequence is an ISM protein described in the exemplification of the invention	
SO	Sequence 1620 AA:	
Query Match	100.0%; Score 52;	DB 5; Length 1620;
Best Local Similarity	100.0%;	Pred. No. 1.7;
Matches 10;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0.
DQ	1 ALPIEXGPLV 10       	
Db	1396 ALPIEXGPLV 1405	
RESULT 14		
ID	ABG74471 standard; protein; 1620 AA.	
XX	ABG74471	
AC	ABG74471;	
DT	11-APR-2003 (first entry)	
DE	Human ALK protein SEQ ID 2.	
XX	ALK; human; nucleophosmin; NPM; anaplastic lymphoma kinase; chromosome 5; t(2); 5) Lymphoma cell; chromosomal rearrangement; chromosome 2;	
KX	nucleolar phosphoprotein; centromeric; telomeric.	
XX	Homo sapiens.	
OS		
XX	Key	Location/Qualifiers
FT	Peptide	1..26
FT	/label=signal_peptide	
FT	Protein	27..1620
FT	/label=mature_ALK	
FT	Domain	1031..1058
FT	/label=transmembrane_domain	
FT	Misc-difference	1058..1059
FT	/note="NPM-ALK fusion junction"	
FT	Domain	1123..1377
FT	/label=tyrosine_kinase_catalytic_domain	
XX	US6451997-B1.	
XX	17-SEP-2002.	

XX 28-SEP-2000; 2000US-00670827.  
 XX 03-DEC-1993; 93US-00160861.  
 XX 12-OCT-1995; 95US-00542363.  
 XX 19-JUN-1998; 98US-00100089.  
 XX (SCUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 XX Morris SW, Look AT.  
 XX WPI; 2003-101739/09.  
 DR N-PSDB; ABQ77106.  
 PT Kit for use in method of detecting t(2;5) chromosomal rearrangements or  
 PT rearrangements involving nucleolar phosphoprotein (NPM) or anaplastic  
 PT lymphoma kinase (ALK) comprises probes that hybridizes with NPM and ALK  
 PT genes.  
 XX  
 XX Example 1; Fig 3B; 87pp; English.  
 XX  
 CC This invention describes a novel kit for use in a method of detecting t(2  
 CC :5) chromosomal rearrangements (CR) or CRs involving nucleolar  
 CC phosphoprotein (NPM) and anaplastic lymphoma kinase (ALK). The method  
 CC comprises labeled probes that hybridize to a sequence of 2442 or 6226  
 CC nucleotides or their complement, and to region of human chromosome 5/2  
 CC having wild-type NPM/ALK genes (the region is centromeric/telomeric to  
 CC nucleotide). This sequence represents a polypeptide sequence described in  
 CC the disclosure of the invention  
 XX  
 SQ Sequence 1620 AA;

Query Match 100.0%; Score 52; DB 6; Length 1620;  
 Best Local Similarity 100.0%; Pred. No. 1.7;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ALPIRYGPLYV 10  
 DB 1396 ALPIRYGPLYV 1405

RESULT 15  
 ADF06379  
 ID ADF06379 standard; protein; 207 AA.  
 XX  
 AC ADF06379;  
 XX  
 DT 12-FEB-2004 (first entry)  
 XX  
 DE Bacterial polypeptide #2492.  
 XX  
 KW Proteus mirabilis infection; bacterial infection; antibacterial;  
 KW immunostimulant.  
 XX  
 OS Proteus mirabilis.  
 XX  
 PN US6605709-B1.  
 XX  
 PD 12-AUG-2003.  
 XX  
 PF 05-APR-2000; 2000US-00543681.  
 XX  
 PR 09-APR-1999; 99US-0128706P.  
 XX  
 PA (GENO-) GENOME THERAPEUTICS CORP.  
 XX  
 PI Breton GL;  
 XX  
 DR WPI; 2003-895291/82.  
 DR N-PSDB; ADF02207.  
 XX  
 PT New Proteus mirabilis polypeptides and polynucleotides, useful as  
 PT reagents for diagnosis of bacterial disease, as components of

PT antibacterial vaccines, as targets for antibacterial drugs, or as  
 PT biocontrol agents for plants.  
 XX  
 PS Disclosure; SEQ ID NO 6664; 870pp; English.  
 XX  
 CC The invention relates to new Proteus mirabilis polypeptides and  
 CC polynucleotides. The invention also relates to antibodies against the  
 CC polynucleotides. The invention also relates to producing the polypeptides, a method of  
 CC generating vaccines for immunizing an individual against P. mirabilis, a  
 CC method for evaluating a compound for the ability to bind a P. mirabilis  
 CC polypeptide and a method for screening test compounds for anti-bacterial  
 CC activity. The polypeptides and polynucleotides are useful as molecular  
 CC targets for diagnosing, preventing and treating pathological conditions  
 CC resulting from bacterial infection, as reagents for diagnosis of  
 CC bacterial diseases, as components of antibacterial vaccines, as targets  
 CC for antibacterial drugs or as bio-control agents for plants. This  
 CC sequence represents a Proteus mirabilis polypeptide of the invention.  
 XX

Query Match 73.1%; Score 38; DB 7; Length 207;  
 Best Local Similarity 60.0%; Pred. No. 73;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ALPIRYGPLYV 10  
 DB 110 ALDMEYGPII 119

Search completed: November 30, 2004, 07:39:57  
 Job time : 79.4295 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: November 30, 2004, 07:19:22 ; Search time 376.923 Seconds  
(without alignments)  
9,410 Million cell updates/sec

Title: US-10-008-377a-6

Perfect score: 52

Sequence: 1 ALPIRYGPLYV 10

## Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1575965 seqs, 354694765 residues

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*

2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*

3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*

4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*

5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*

6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*

7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*

8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*

9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*

10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*

11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*

12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*

13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*

14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*

15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*

16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*

17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*

18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*

19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*

20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	10	US-10-008-377-6	Sequence 6, Appl1
2	52	100.0	370	US-10-664-421-70	Sequence 70, Appl1
3	52	100.0	680	US-09-827-949-4	Sequence 4, Appl1
4	52	100.0	1620	US-09-827-949-2	Sequence 2, Appl1
5	39	75.0	121	US-10-424-599-264129	Sequence 264129, A
6	39	75.0	685	US-10-425-114-45829	Sequence 586, App
7	38	73.1	401	US-10-389-647-586	Sequence 34, Appl1
8	37	71.2	58	US-10-038-722-34	Sequence 51, Appl1
9	37	71.2	60	US-10-456-868-51	Sequence 55, Appl1
10	37	71.2	206	US-10-211-962-55	Sequence 45922, A
11	37	71.2	677	US-10-425-114-45922	Sequence 330, App
12	37	71.2	1083	US-10-080-170-330	Sequence 330, App
13	37	71.2	1083	US-10-080-170-330	Sequence 330, App

14	37	71.2	1083	US-10-468-356-330	Sequence 2330, App
15	36	69.2	287	US-10-424-599-213533	Sequence 213533, App
16	36	69.2	446	US-10-425-115-267156	Sequence 267156, App
17	36	69.2	695	US-10-437-963-180466	Sequence 180466, App
18	36	69.2	728	US-10-282-122A-46079	Sequence 46079, A
19	35	67.3	51	US-09-827-948-7	Sequence 7, Appl1
20	35	67.3	51	US-09-741-106-2	Sequence 2, Appl1
21	35	67.3	51	US-10-176-071-7	Sequence 15, Appl1
22	35	67.3	55	US-09-904-621-15	Sequence 21, Appl1
23	35	67.3	57	US-09-974-026-21	Sequence 70, Appl1
24	35	67.3	58	US-10-167-351-70	Sequence 37, Appl1
25	35	67.3	58	US-10-038-722-32	Sequence 12, Appl1
26	35	67.3	58	US-10-456-986A-37	Sequence 11, Appl1
27	35	67.3	58	US-10-361-997-12	Sequence 51101, A
28	35	67.3	61	US-10-767-701-51101	Sequence 71, Appl1
29	35	67.3	78	US-10-799-326-71	Sequence 26, Appl1
30	35	67.3	88	US-09-978-418-26	Sequence 1, Appl1
31	35	67.3	123	US-10-356-088-1	Sequence 21, Appl1
32	35	67.3	123	US-10-356-088-23	Sequence 1, Appl1
33	35	67.3	123	US-10-353-454-1	Sequence 23, Appl1
34	35	67.3	123	US-10-353-454-23	Sequence 1, Appl1
35	35	67.3	123	US-10-799-326-1	Sequence 1, Appl1
36	35	67.3	123	US-10-799-326-23	Sequence 2, Appl1
37	35	67.3	124	US-10-422-536-10	Sequence 10, Appl1
38	35	67.3	213	US-09-766-778-2	Sequence 6, Appl1
39	35	67.3	213	US-10-086-176A-6	Sequence 2, Appl1
40	35	67.3	235	US-09-904-621-2	Sequence 332, App
41	35	67.3	235	US-09-736-457-332	Sequence 332, App
42	35	67.3	235	US-09-902-941-332	Sequence 332, App
43	35	67.3	235	US-09-849-626-332	Sequence 332, App
44	35	67.3	235	US-09-476-300-332	Sequence 332, App
45	35	67.3	235	US-10-017-754-332	Sequence 332, App

## ALIGNMENTS

RESULT 1

US-10-008-377-6

Sequence 6, Application US/10008377

Publication No. US20030157101A1

GENERAL INFORMATION:

APPLICANT: Gambacorti-Passerini, Carlo

APPLICANT: Passoni, Lorena

TITLE OF INVENTION: Immunogenic ALK Peptides

FILE REFERENCE: 045922/241203

CURRENT APPLICATION NUMBER: US/10/008,377

CURRENT FILING DATE: 2001-11-15

NUMBER OF SEQ ID NOS: 7

SOFTWARE: FastSeq For Windows Version 4.0

SEQ ID NO 6

LENGTH: 10

TYPE: PRT

ORGANISM: Homo sapiens

US-10-008-377-6

Query Match

Best Local Similarity 100.0%; Score 52; DB 14; Length 10;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 ALPIRYGPLYV 10

DB

1 ALPIRYGPLYV 10

RESULT 2

US-10-664-421-70

Sequence 70, Application US/10664421

Publication No. US20040142864A1

GENERAL INFORMATION:

APPLICANT: BREWER, RYAN

APPLICANT: IBRAHIM, PRABHA

APPLICANT: KUMAR, ABHINAV

```

; APPLICANT: MANDIYAN, VALISAN
; APPLICANT: MILBURN, MICHAEL V.
; TITLE OF INVENTION: CRYSTAL STRUCTURE OF P1M-1 KINASE
; FILE REFERENCE: 039363/0703
; CURRENT APPLICATION NUMBER: US/10/664,421
; CURRENT FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 60/412,341
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/411,398
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 70
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-664-421-70

Query Match      100.0%; Score 52; DB 16; Length 370;
Best Local Similarity 100.0%; Pred. No. 0.42;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ALPIEYGPLV 10
DB      331 ALPIEYGPLV 340

RESULT 3
US-09-827-949-4
; Sequence 4, Application US/09827949
; Patent No. US20010021505A1
; GENERAL INFORMATION:
; APPLICANT: MORRIS, STEPHAN W.
; TITLE OF INVENTION: ALK PKCprotein Tyrosine Kinase/Receptor and Ligands Thereof
; FILE REFERENCE: 0656.040004
; CURRENT APPLICATION NUMBER: US/09/827,949
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 09/670,827
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 09/100,089
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: US 08/542,363
; PRIOR FILING DATE: 1995-10-12
; PRIOR APPLICATION NUMBER: US 08/160,861
; PRIOR FILING DATE: 1993-12-03
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 680
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-827-949-4

Query Match      100.0%; Score 52; DB 9; Length 680;
Best Local Similarity 100.0%; Pred. No. 0.79;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ALPIEYGPLV 10
DB      456 ALPIEYGPLV 465

RESULT 4
US-09-827-949-2
; Sequence 2, Application US/09827949
; Patent No. US20010021505A1
; GENERAL INFORMATION:
; APPLICANT: MORRIS, STEPHAN W.
; APPLICANT: LOCK, A. THOMAS
; TITLE OF INVENTION: ALK PKCprotein Tyrosine Kinase/Receptor and Ligands Thereof
; FILE REFERENCE: 0656.040004
; CURRENT APPLICATION NUMBER: US/09/827,949
```

```

; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 09/670,827
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 09/100,089
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: US 08/542,363
; PRIOR FILING DATE: 1995-10-12
; PRIOR APPLICATION NUMBER: US 08/160,861
; PRIOR FILING DATE: 1993-12-03
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1620
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-827-949-2

Query Match      100.0%; Score 52; DB 9; Length 1620;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ALPIEYGPLV 10
DB      1396 ALPIEYGPLV 1405

RESULT 5
US-10-424-599-264129
; Sequence 264129, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: LA ROSA THOMAS J
; APPLICANT: KOVALIC DAVID K
; APPLICANT: ZHOU YIHUA
; APPLICANT: CAO YONGWEI
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53123)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 265684
; SEQ ID NO 264129
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_80529C.1.pep
US-10-424-599-264129

Query Match      75.0%; Score 39; DB 15; Length 121;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      2 LPIEYGPLV 10
DB      94 LPIEYGPLV 102

RESULT 6
US-10-425-114-45829
; Sequence 45829, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: LIU, JINGDONG
; APPLICANT: ZHOU, YIHUA
; APPLICANT: KOVALIC, DAVID K.
; APPLICANT: SCREEN, STEVEN E
; APPLICANT: TABASKA, JACK E
; APPLICANT: CAO, YONGWEI
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
```

;; CURRENT FILING DATE: 2003-04-28  
;; NUMBER OF SEQ ID NOS: 73128  
;; SEQ ID NO 45829  
;; LENGTH: 685  
;; TYPE: PRT  
;; ORGANISM: Glycine max  
;; FEATURES:  
;; OTHER INFORMATION: Clone ID: 700910011\_Flt.pcp  
US-10-425-114-45829

Query Match 75.0%; Score 39; DB 15; Length 685;  
Best Local Similarity 60.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALPIRYGPLY 10  
DB 147 SLPLKXGPFV 156

RESULT 7  
US-10-389-647-586  
;; Sequence 586, Application US/10389647  
;; Publication No. US20040033549A1  
;; GENERAL INFORMATION:  
;; APPLICANT: GREENBERG, E. Peter  
;; APPLICANT: SCHUSTER, Martin  
;; APPLICANT: LOSTROH, Carol  
;; TITLE OF INVENTION: QUORUM SENSING SIGNALING IN BACTERIA  
;; FILE REFERENCE: UIZ-038CP  
;; CURRENT APPLICATION NUMBER: US/10/389,647  
;; CURRENT FILING DATE: 2003-03-14  
;; PRIOR APPLICATION NUMBER: 09/653730  
;; PRIOR FILING DATE: 2000-09-01  
;; PRIOR APPLICATION NUMBER: 60/153022  
;; PRIOR FILING DATE: 1999-09-03  
;; NUMBER OF SEQ ID NOS: 710  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 586  
;; LENGTH: 401  
;; TYPE: PRT  
;; ORGANISM: Pseudomonas aeruginosa  
US-10-389-647-586

Query Match 73.1%; Score 38; DB 15; Length 401;  
Best Local Similarity 88.9%; Pred. No. 1.6e+02;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALPIRYGPLY 9  
DB 278 ALPIRYGPLY 286

RESULT 8  
US-10-038-722-34  
;; Sequence 34, Application US/10038722  
;; Publication No. US20030175919A1  
;; GENERAL INFORMATION:  
;; APPLICANT: LEY, Arthur C.  
;; APPLICANT: GUTERMAN, Sonia K.  
;; APPLICANT: MARKLAND, William  
;; APPLICANT: KENT, Rachel B.  
;; APPLICANT: ROBERTS, Bruce L.  
;; APPLICANT: LADNER, Robert C.  
;; TITLE OF INVENTION: ITI-D1 KUNITZ DOMAIN MUTANTS AS THE INHIBITORS  
;; FILE REFERENCE: LEY-1B  
;; CURRENT APPLICATION NUMBER: US/10/038,722  
;; CURRENT FILING DATE: 2002-01-08  
;; PRIOR APPLICATION NUMBER: US 08/649,406  
;; PRIOR FILING DATE: 1999-07-21  
;; PRIOR APPLICATION NUMBER: PCT/US95/16349  
;; PRIOR FILING DATE: 1995-12-15  
;; PRIOR APPLICATION NUMBER: US 08/358,160  
;; PRIOR FILING DATE: 1994-12-16

;; NUMBER OF SEQ ID NOS: 129  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 34  
;; LENGTH: 58  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURES:  
;; OTHER INFORMATION: DPL 2.2  
US-10-038-722-34

Query Match 71.2%; Score 37; DB 14; Length 58;  
Best Local Similarity 55.6%; Pred. No. 32;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 LPRIYGPV 10  
DB 7 LPDIYGPV 15

RESULT 9  
US-10-456-986A-51  
;; Sequence 51, Application US/10456986A  
;; Publication No. US20040038931A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Ladner, Robert C.  
;; APPLICANT: Ley, Arthur C.  
;; APPLICANT: HIRANI, Shirish  
;; APPLICANT: Williams, Anthony  
;; TITLE OF INVENTION: Prevention and Reduction of Blood Loss  
;; FILE REFERENCE: 3421.1001-002  
;; CURRENT APPLICATION NUMBER: US/10/456,986A  
;; CURRENT FILING DATE: 2003-06-06  
;; PRIOR APPLICATION NUMBER: 60/387,239  
;; PRIOR FILING DATE: 2002-06-07  
;; PRIOR APPLICATION NUMBER: 60/407,003  
;; PRIOR FILING DATE: 2002-08-28  
;; NUMBER OF SEQ ID NOS: 54  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 51  
;; LENGTH: 60  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURES:  
;; OTHER INFORMATION: DPL24 KR Sequence  
US-10-456-986A-51

Query Match 71.2%; Score 37; DB 15; Length 60;  
Best Local Similarity 55.6%; Pred. No. 32;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 LPRIYGPV 10  
DB 9 LPDIYGPV 17

RESULT 10  
US-10-211-962-55  
;; Sequence 55, Application US/10211962  
;; Publication No. US20030082640A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Herz, Joachim  
;; APPLICANT: Gotthardt, Michael  
;; TITLE OF INVENTION: LDL Receptor Signaling Pathways  
;; FILE REFERENCE: UTSW0708  
;; CURRENT APPLICATION NUMBER: US/10/211,962  
;; CURRENT FILING DATE: 2002-08-01  
;; PRIOR APPLICATION NUMBER: US/09/562,737  
;; PRIOR FILING DATE: 2000-05-01  
;; NUMBER OF SEQ ID NOS: 132  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 55  
;; LENGTH: 206  
;; TYPE: PRT

ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
OTHER INFORMATION: Sequence  
US-10-211-962-55

Query Match 71.2%; Score 37; DB 14; Length 206;  
Best Local Similarity 55.6%; Pred. No. 1.2e+02;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPPIYGPV 10  
DB 161 LPVQMGPIV 169

RESULT 11  
US-10-425-114-45922  
Sequence 45922, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jingdong  
APPLICANT: Zhou, Yinhua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E  
APPLICANT: Tabaska, Jack E  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21133131B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 45922  
LENGTH: 677  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: 701011172\_FLI.pep  
US-10-425-114-45922

Query Match 71.2%; Score 37; DB 15; Length 677;  
Best Local Similarity 55.6%; Pred. No. 4.1e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 LPPIYGPV 10  
DB 142 LPLKXGPI 150

RESULT 12  
US-10-080-170-330  
Sequence 330, Application US/10080170  
Publication No. US20030129601A1  
GENERAL INFORMATION:  
APPLICANT: COLE, S.T.  
TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR  
TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR  
FILE REFERENCE: 03495,0218  
CURRENT APPLICATION NUMBER: US/10/080,170  
CURRENT FILING DATE: 2002-06-10  
PRIOR APPLICATION NUMBER: 60/270,123  
PRIOR FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 652  
SOFTWARE: Patent Ver. 2.1  
SEQ ID NO 330  
LENGTH: 1083  
TYPE: PRT  
ORGANISM: Mycobacterium leprae  
US-10-080-170-330

Query Match 71.2%; Score 37; DB 14; Length 1083;  
Best Local Similarity 77.8%; Pred. No. 6.7e+02;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ALPIEYGPL 9  
DB 753 ALPSNYGPL 761

RESULT 13  
US-10-080-170-330  
Sequence 330, Application US/10080170  
Publication No. US20040121322A9  
GENERAL INFORMATION:  
APPLICANT: COLE, S.T.  
TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR  
TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR  
FILE REFERENCE: 03495,0218  
CURRENT APPLICATION NUMBER: US/10/080,170  
CURRENT FILING DATE: 2002-06-10  
PRIOR APPLICATION NUMBER: 60/270,123  
PRIOR FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 652  
SOFTWARE: Patent Ver. 2.1  
SEQ ID NO 330  
LENGTH: 1083  
TYPE: PRT  
ORGANISM: Mycobacterium leprae  
US-10-080-170-330

Query Match 71.2%; Score 37; DB 16; Length 1083;  
Best Local Similarity 77.8%; Pred. No. 6.7e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ALPIEYGPL 9  
DB 753 ALPSNYGPL 761

RESULT 14  
US-10-468-356-330  
Sequence 330, Application US/10468356  
Publication No. US20040197896A1  
GENERAL INFORMATION:  
APPLICANT: COLE, STEWART  
TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR  
TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR  
FILE REFERENCE: 03394,0019  
CURRENT APPLICATION NUMBER: US/10/468,356  
CURRENT FILING DATE: 2003-08-19  
PRIOR APPLICATION NUMBER: 10/080,170  
PRIOR FILING DATE: 2002-02-22  
PRIOR APPLICATION NUMBER: 60/270,123  
PRIOR FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 655  
SOFTWARE: Patent Ver. 3.2  
SEQ ID NO 330  
LENGTH: 1083  
TYPE: PRT  
ORGANISM: Mycobacterium leprae  
US-10-468-356-330

Query Match 71.2%; Score 37; DB 17; Length 1083;  
Best Local Similarity 77.8%; Pred. No. 6.7e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ALPIEYGPL 9  
DB 753 ALPSNYGPL 761

RESULT 15  
US-10-424-599-21353

Tue Nov 30 08:49:27 2004

us-10-008-377a-6.open.rapb

Page 5

```
; Sequence 213533, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5323)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 213533
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_34845C.1.pep
US-10-424-599-213533

Query Match          69.2%; Score 36; DB 15; Length 287;
Best Local Similarity 70.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

CY      1 ALPLEYGLY 10
      |||||
Db      56 ALPEYDPIV 65

Search completed: November 30, 2004, 08:06:36
Job time : 377.923 secs
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 / Search time 25.1282 Seconds  
(without alignments)  
26.392 Million cell updates/sec

Title: US-10-008-377A-6

Sequence: 1 ALPIEXGPLV 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents AA:  
1: /cgn2\_6/prodata/1/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/prodata/1/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/prodata/1/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/prodata/1/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/prodata/1/1aa/PCTUS.COMB.pep.\*  
6: /cgn2\_6/prodata/1/1aa/backfilest.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	100.0	525	1	US-08-160-861-4
2	52	100.0	572	1	US-08-160-861-3
3	52	100.0	680	1	US-08-542-363-4
4	52	100.0	680	3	US-09-100-089-4
5	52	100.0	680	3	US-09-100-089-4
6	52	100.0	680	4	US-09-670-827-4
7	52	100.0	680	4	US-09-827-949-4
8	52	100.0	1620	3	US-08-542-363-2
9	52	100.0	1620	3	US-09-100-089-2
10	52	100.0	1620	4	US-09-670-827-2
11	52	100.0	1620	4	US-09-827-949-2
12	38	73.1	207	4	US-09-543-681A-6664
13	38	73.1	409	4	US-09-252-991A-18004
14	37	71.2	206	4	US-09-562-737-55
15	36	69.2	243	4	US-09-252-991A-69268
16	36	69.2	393	4	US-09-270-767-38581
17	35	67.3	393	4	US-09-270-767-53798
18	35	67.3	51	1	US-08-437-841-2
19	35	67.3	51	1	US-08-286-521-2
20	35	67.3	51	1	US-08-436-175-2
21	35	67.3	51	3	US-08-943-682-2
22	35	67.3	51	3	US-09-913-896A-7
23	35	67.3	51	4	US-09-827-948-7
24	35	67.3	51	4	US-09-741-106-2
25	35	67.3	51	5	PCT-US95-09464-2
26	35	67.3	55	1	US-08-147-710-15
27	35	67.3	55	1	US-08-458-090-15
					Sequence 15, App

28	35	67.3	55	4	US-09-904-621-15	Sequence 15, App
29	35	67.3	57	4	US-09-144-428-21	Sequence 21, App
30	35	67.3	58	1	US-08-358-160-80	Sequence 80, App
31	35	67.3	58	1	US-08-676-125A-39	Sequence 39, App
32	35	67.3	58	2	US-09-136-012A-39	Sequence 70, App
33	35	67.3	58	3	US-08-676-124-70	Sequence 70, App
34	35	67.3	58	3	US-09-414-878-70	Sequence 70, App
35	35	67.3	58	3	US-09-240-136-70	Sequence 70, App
36	35	67.3	58	4	US-09-638-770A-70	Sequence 70, App
37	35	67.3	124	2	US-09-800-170-10	Sequence 10, App
38	35	67.3	213	2	US-08-796-850-2	Sequence 2, App
39	35	67.3	213	4	US-08-766-778-2	Sequence 2, App
40	35	67.3	235	1	US-08-147-710-2	Sequence 2, App
41	35	67.3	235	1	US-08-458-090-2	Sequence 2, App
42	35	67.3	235	2	US-08-457-887-2	Sequence 2, App
43	35	67.3	235	3	US-08-817-145-3	Sequence 3, App
44	35	67.3	235	4	US-09-702-705-332	Sequence 332, App
45	35	67.3	235	4	US-09-736-457-332	Sequence 332, App

#### ALIGNMENTS

RESULT 1  
US-08-160-861-4  
; Sequence 4, Application US/08160861  
; Patent No. 5529295  
; APPLICANT: MORRIS, STEPHAN W  
; TITLE OF INVENTION: NOVEL FUSION NUCLEIC ACID SEQUENCES AND  
; TITLE OF INVENTION: FUSION PROTEINS PRESENT IN HUMAN T(2:5) LYMPHOMA, METHODS  
; TITLE OF INVENTION: OF DETECTION AND USES THEREOF  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX  
; STREET: 1100 NEW YORK AVE NW SUITE 600  
; CITY: WASHINGTON  
; STATE: D.C.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/160,861  
; FILING DATE: 02-DEC-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MILLMAN, ROBERT A  
; REGISTRATION NUMBER: 36217  
; REFERENCE/DOCKET NUMBER: 0656.0400000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2678  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 525 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-160-861-4  
Query Match 100.0%; Score 52; DB 1; Length 525;  
Best Local Similarity 100.0%; Pred. No. 0.2;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 ALPIEXGPLV 10  
DB 456 ALPIEXGPLV 465  
RESULT 2  
US-08-160-861-3

```
Sequence 3, Application US/08160861
Patent No. 5529295
GENERAL INFORMATION:
APPLICANT: MORRIS, STEPHAN W
TITLE OF INVENTION: NOVEL FUSION NUCLEIC ACID SEQUENCES AND
TITLE OF INVENTION: FUSION PROTEINS PRESENT IN HUMAN T(2:5) LYMPHOMA, METHODS
OF DETECTION AND USE THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX
STREET: 1100 NEW YORK AVE NW SUITE 600
CITY: WASHINGTON
STATE: D.C.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/160,861
FILING DATE: 02-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MILLMAN, ROBERT A
REGISTRATION NUMBER: 36217
REFERENCE/DOCKET NUMBER: 0656.0400000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2678
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 572 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-160-861-3

Query Match
Best Local Similarity 100.0%; Score 52; DB 1; Length 572;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIRYGPLY 10
DB 503 ALPIRYGPLY 512

RESULT 3
US-08-542-363-4
Sequence 4, Application US/08542363
Patent No. 5770421
GENERAL INFORMATION:
APPLICANT: MORRIS, STEPHAN W.
APPLICANT: LOOK, A. THOMAS
TITLE OF INVENTION: ALK PROTEIN TYROSINE KINASE/RECEPTOR AND
TITLE OF INVENTION: LIGANDS THEREOF
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/542,363
FILING DATE: 12-OCT-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
```

```
NAME: FOX, SAMUEL L.
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 0656.0400001/SLE/GKT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 680 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-542-363-4

Query Match
Best Local Similarity 100.0%; Score 52; DB 1; Length 680;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIRYGPLY 10
DB 456 ALPIRYGPLY 465

RESULT 4
US-09-100-089-4
Sequence 4, Application US/09100089
Patent No. 6174674
GENERAL INFORMATION:
APPLICANT: MORRIS, STEPHAN W.
APPLICANT: LOOK, A. THOMAS
TITLE OF INVENTION: ALK PROTEIN TYROSINE KINASE/RECEPTOR AND
TITLE OF INVENTION: LIGANDS THEREOF
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,089
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/160,861
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/542,363
FILING DATE: 12-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: FOX, SAMUEL L.
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 0656.0400002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 680 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-100-089-4

Query Match
Best Local Similarity 100.0%; Score 52; DB 3; Length 680;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 ALPIEYGLV 10  
DB 456 ALPIEYGLV 465

RESULT 5  
US-09-670-827-4

Sequence 4, Application US/09670827  
Patent No. 6431997

GENERAL INFORMATION:

APPLICANT: Morris, Stephan W.  
Look, A. Thomas

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
Ligands Thereof

NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:

ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600

CITY: Washington  
STATE: DC

COUNTRY: USA  
ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/670,827  
FILING DATE: 28-Sep-2000

CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/160,861  
FILING DATE: 03-Dec-1993

APPLICATION NUMBER: US 08/542,363  
FILING DATE: 12-Oct-1995

APPLICATION NUMBER: US 09/100,089  
FILING DATE: 19-Jun-1998

ATTORNEY/AGENT INFORMATION:

NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353

REFERENCE/DOCKET NUMBER: 0656.0400003  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 680 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-670-827-4

Query Match 100.0%; Score 52; DB 4; Length 680;  
Best Local Similarity 100.0%; Pred. No. 0.27;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIEYGLV 10  
DB 456 ALPIEYGLV 465

RESULT 6  
US-09-827-949-4

Sequence 4, Application US/09827949  
Patent No. 6636548

GENERAL INFORMATION:

APPLICANT: Morris, Stephan W.  
Look, A. Thomas

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof  
FILE REFERENCE: 0656.0400004

CURRENT APPLICATION NUMBER: US/09/827,949  
CURRENT FILING DATE: 2001-04-09  
PRIOR APPLICATION NUMBER: US 09/670,827  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 09/100,089  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: US 08/542,363  
PRIOR FILING DATE: 1995-10-12  
PRIOR APPLICATION NUMBER: US 08/160,861  
PRIOR FILING DATE: 1993-12-03  
NUMBER OF SEQ ID NOS: 43  
SOFTWARE: Patent version 3.0  
SEQ ID NO 4  
LENGTH: 680  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-827-949-4

Query Match 100.0%; Score 52; DB 4; Length 680;  
Best Local Similarity 100.0%; Pred. No. 0.27;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIEYGLV 10  
DB 456 ALPIEYGLV 465

RESULT 7  
US-08-542-363-2

Sequence 2, Application US/08542363  
Patent No. 5770421

GENERAL INFORMATION:

APPLICANT: Morris, Stephan W.  
Look, A. Thomas

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
Ligands Thereof

NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:

ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600

CITY: Washington  
STATE: DC

COUNTRY: USA  
ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/542,363  
FILING DATE: 12-Oct-1995

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:

NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353

REFERENCE/DOCKET NUMBER: 0656.0400001/SLF/GKT  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1620 amino acids  
TYPE: amino acid

STRANDEDNESS:  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-542-363-2

Query Match 100.0%; Score 52; DB 1; Length 1620;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 ALPIEYGPLV 10  
|||  
Db 1396 ALPIEYGPLV 1405

RESULT 8  
US-09-100-089-2

Sequence 2, Application US/09100089

Patent No. 6174674

GENERAL INFORMATION:

APPLICANT: Morris, Stephan W.

APPLICANT: Look, A. Thomas

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.

STREET: 1100 New York Avenue, N.W., Suite 600

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/100,089

FILING DATE: Herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/160,861

FILING DATE: 03-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/542,363

FILING DATE: 12-OCT-1995

ATTORNEY/AGENT INFORMATION:

NAME: Fox, Samuel L.

REGISTRATION NUMBER: 30,353

REFERENCE/DOCKET NUMBER: 0656.0400002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1620 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-100-089-2

Query Match 100.0%; Score 52; DB 3; Length 1620;

Best Local Similarity 100.0%; Pred. No. 0.67;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 ALPIEYGPLV 10  
|||  
Db 1396 ALPIEYGPLV 1405

RESULT 9  
US-09-670-827-2

Sequence 2, Application US/09670827

Patent No. 6451957

GENERAL INFORMATION:

APPLICANT: Morris, Stephan W.

APPLICANT: Look, A. Thomas

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/670,827

FILING DATE: 28-Sep-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/160,861

FILING DATE: 03-DEC-1993

APPLICATION NUMBER: US 08/542,363

FILING DATE: 12-OCT-1995

APPLICATION NUMBER: US 09/100,089

FILING DATE: 19-JUN-1998

ATTORNEY/AGENT INFORMATION:

NAME: Fox, Samuel L.

REGISTRATION NUMBER: 30,353

REFERENCE/DOCKET NUMBER: 0656.0400003

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1620 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-670-827-2

Query Match 100.0%; Score 52; DB 4; Length 1620;

Best Local Similarity 100.0%; Pred. No. 0.67; Mismatches 0; Indels 0; Gaps 0;

CY 1 ALPIEYGPLV 10  
|||  
Db 1396 ALPIEYGPLV 1405

RESULT 10  
US-09-827-949-2

Sequence 2, Application US/09827949

Patent No. 6696548

GENERAL INFORMATION:

APPLICANT: Morris, Stephan W.

APPLICANT: Look, A. Thomas

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.

STREET: 1100 New York Avenue, N.W., Suite 600

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/827,949

FILING DATE: 2001-04-09

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/670,827

FILING DATE: 2000-09-28

PRIOR APPLICATION NUMBER: US 09/100,089

FILING DATE: 1998-06-19

PRIOR APPLICATION NUMBER: US 08/542,363

FILING DATE: 1995-10-12

PRIOR APPLICATION NUMBER: US 08/160,861

FILING DATE: 1993-12-03

NUMBER OF SEQ ID NOS: 43

SOFTWARE: Patentin version 3.0

SEQ ID NO: 2

LENGTH: 1620

TYPE: PRT

ORGANISM: Homo sapiens

US-09-827-949-2

Query Match 100.0%; Score 52; DB 4; Length 1620;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALPIEYGPLV 10  
Db 1396 ALPIEYGPLV 1405

RESULT 11  
US-09-543-681A-6664  
; Sequence 6664; Application US/09543681A  
; Patent No. 6605708  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 6664  
; LENGTH: 207  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-6664

Query Match 73.1%; Score 39; DB 4; Length 207;  
Best Local Similarity 60.0%; Pred. No. 24;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALPIEYGPLV 10  
Db 110 ALDMEYGPIL 119

RESULT 12  
US-09-252-991A-18004  
; Sequence 18004; Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: MARC J. RUBENFELD et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 18004  
; LENGTH: 409  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-18004

Query Match 73.1%; Score 38; DB 4; Length 409;  
Best Local Similarity 88.9%; Pred. No. 49;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALPIEYGPL 9  
Db 286 ALPIEYGPL 294

RESULT 13  
US-09-562-737-55

; Sequence 55; Application US/09562737

; Patent No. 6428967  
; GENERAL INFORMATION:  
; APPLICANT: Herz, Joachim  
; APPLICANT: Gottardt, Michael  
; TITLE OF INVENTION: Ldb Receptor Signaling Pathways  
; FILE REFERENCE: UTSW0708  
; CURRENT APPLICATION NUMBER: US/09/562,737  
; CURRENT FILING DATE: 2000-05-01  
; NUMBER OF SEQ ID NOS: 132  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 55  
; LENGTH: 206  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-562-737-55

Query Match 71.2%; Score 37; DB 4; Length 206;  
Best Local Similarity 55.6%; Pred. No. 36;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LPVQWGPV 10  
Db 161 LPVQWGPV 169

RESULT 14  
US-09-252-991A-29268  
; Sequence 29268; Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: MARC J. RUBENFELD et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 29268  
; LENGTH: 243  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-29268

Query Match 69.2%; Score 36; DB 4; Length 243;  
Best Local Similarity 75.0%; Pred. No. 65;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALPIEYGP 8  
Db 206 ALPIEYGP 213

RESULT 15  
US-09-270-767-38581  
; Sequence 38581; Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 38581

LENGTH: 393  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
FEATURE:  
OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-38581

Query Match 69.2%; Score 36; DB 4; Length 393;  
Best Local Similarity 85.7%; Pred. No. 1.1e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

CY 3 PIRXGPI 9  
Db 107 PIRGPI 113

Search completed: November 30, 2004, 07:19:16  
Job time : 26.1282 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 ; Search time 18 Seconds  
(without alignments)  
48.108 Million cell updates/sec

Title: US-10-008-377a-7  
Perfect score: 42  
Sequence: 1 SLTNMKEV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: PIR\_79:\*

1: p1r1:\*

2: p1r2:\*

3: p1r3:\*

4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	42	100.0	680 4	138491 nucleophosmin/anap
2	33	78.6	206 2	122345 hypothetical prote
3	32	76.2	341 2	A11186 transport protei
4	31	73.8	510 2	A11196 D-alanine-activa
5	31	73.8	510 2	AD1554 D-alanine-activa
6	31	73.8	544 2	AH1544 transport protei
7	31	73.8	583 2	S38642 chaperonin 60 alp
8	31	73.8	599 2	F82291 conserved hypotet
9	30	71.4	145 2	A1259 transcription regu
10	30	71.4	232 2	C95196 conserved hypotet
11	30	71.4	232 2	H98062 conserved hypotet
12	30	71.4	232 2	C98021 N-acetylmannosamin
13	30	71.4	233 2	C95154 N-acetylmannosamin
14	30	71.4	285 2	A86732 geranyltransferas
15	30	71.4	291 2	A6498 chemotactic methyl
16	30	71.4	340 2	S20879 homeotic protein H
17	30	71.4	347 2	AF2645 flagellar motor sw
18	30	71.4	347 2	F97427 flagellar motor sw
19	30	71.4	550 1	FGRTA fibrinogen alpha c
20	30	71.4	551 2	S52287 fibrinogen alpha c
21	30	71.4	608 2	A72213 uridin - alkylom
22	30	71.4	1734 2	A54602 1-deoxyxylulose-5-
23	29	69.0	184 2	SS5435 microtubule-associ
24	29	69.0	198 2	SS5435 hypothetical prote
25	29	69.0	215 2	E75058 conserved hypotet
26	29	69.0	221 2	A11660 hypothetical prote
27	29	69.0	222 2	D65882 hypothetical prote
28	29	69.0	222 2	E72041 uridine kinase CPO
29	29	69.0	222 2	E72041 uridine kinase CPO
29	29	69.0	222 2	AC1289 hypothetical prote

30	29	69.0	243 2	D82681 protein transferas
31	29	69.0	359 2	H90301 hypothetical prote
32	29	69.0	369 2	AB1537 hypothetical prote
33	29	69.0	395 2	T40102 conserved hypotet
34	29	69.0	398 2	F64456 hypothetical prote
35	29	69.0	418 2	A53120 intracellular coag
36	29	69.0	482 2	T39455 adenylosuccinate l
37	29	69.0	482 2	S51377 probable membrane
38	29	69.0	521 2	C28529 nicotinic acetylch
39	29	69.0	611 2	T45493 glutamine-fructose
40	29	69.0	625 2	T25373 hypothetical prote
41	29	69.0	677 2	S15667 transcription fact
42	29	69.0	701 2	S17196 transcription fact
43	29	69.0	727 2	S18193 autocatalygen NOR-90
44	29	69.0	727 2	JC5113 ribosomal transcri
45	29	69.0	727 2	B40439 UBF transcription

#### ALIGNMENTS

##### RESULT 1

138491 nucleophosmin/anaplastic lymphoma kinase mutant fusion protein - human  
C:Species: Homo sapiens (man)  
C:Date: 15-Feb-1996 #sequence\_revision 15-Feb-1996 #text\_change 20-Apr-2000  
C:Accession: 138491  
R:Morris, S.W.; Kirstein, M.N.; Valentine, M.B.; Dittmer, K.G.; Shapiro, D.N.; Saltman, Science 263, 1281-1284, 1994  
A>Title: Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin  
A:Reference number: A53182; MUID:94167588; PMID:8122112  
A:Accession: 138491  
A:Molecule type: mRNA  
A:Residues: 1-680 <NOR>  
A:Cross-references: EMBL:U04946; NID:G609341; PIDN:AA58698.1; PID:G609342  
C:Comment: This sequence is the chimeric product of a translocation mutation.  
C:Genetics:  
A:Gene: NPM1/ALK  
A:Map position: 5/2p23-2p23  
C:Keywords: fusion protein

Query Match 100.0%; Score 42; DB 4; Length 680;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTNMKEV 9  
DB 621 SLTNMKEV 629

##### RESULT 2

T2345 hypothetical protein F47B8.8 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T2345  
R:Berkis, M.; McMurtry, A.  
A:Submitted to the EMBL Data Library, July 1996  
A:Reference number: Z19551  
A:Accession: T2345  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-206 <NID>  
A:Cross-references: UNIPROT:Q20521; EMBL:Z77662; PIDN:CA801196.1; GSPDB:GN00023; CESP:F  
A:Experimental source: Clone F47B8  
C:Genetics:  
A:Gene: CESP:F47B8.8  
A:Map position: 5  
A:introns: 71/1, 171/3

Query Match 78.6%; Score 33; DB 2; Length 206;  
Best Local Similarity 75.0%; Pred. No. 10;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 8  
 Db 131 SLTANKEV 138

## RESULT 3

transport protein homolog lmo0897 [imported] - *Listeria monocytogenes* (strain EGD-e)  
 C/Species: *Listeria monocytogenes*  
 C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C/Accession: AF1186  
 R/Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
 D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A/Authors: Kretz, U.; Kuhn, M.; Kunst, F.; Kurapkak, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schlueker, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.  
 A/Title: Comparative genomics of *Listeria* species.  
 A/Reference number: AB1077; MUID:21537279; PMID:11679669  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-541 <G>  
 A/Cross-references: UNIPROT:Q8Y8K4; GB:NC\_003210; PIDN:CAC98975.1; PID:G16410300; GSPDB: A/Experimental source: strain EGD-e  
 C/Genetics:  
 A/Gene: lmo0897  
 C/Superfamily: integral membrane protein HP0228

Query Match 76.8%; Score 32; DB 2; Length 541;  
 Best Local Similarity 66.7%; Pred. No. 50;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
 Db 154 SFVANKKEI 162

## RESULT 4

D-alanine-activating enzyme (dae), D-alanine-D-alanyl carrier protein ligase (dcl) [imported]  
 C/Species: *Listeria monocytogenes*  
 C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C/Accession: AF1196  
 R/Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
 D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A/Authors: Kretz, U.; Kuhn, M.; Kunst, F.; Kurapkak, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schlueker, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.  
 A/Title: Comparative genomics of *Listeria* species.  
 A/Reference number: AB1077; MUID:21537279; PMID:11679669  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-510 <G>  
 A/Cross-references: UNIPROT:Q8Y8D4; GB:NC\_003210; PIDN:CAC99052.1; PID:G16410376; GSPDB: A/Experimental source: strain EGD-e  
 C/Genetics:  
 A/Gene: dclA  
 C/Superfamily: acetate-CoA ligase; acetate-CoA ligase homology

Query Match 73.8%; Score 31; DB 2; Length 510;  
 Best Local Similarity 55.6%; Pred. No. 78;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
 Db 224 TITANMKDL 232

## RESULT 5

AD1554  
 D-alanine-activating enzyme (dae), D-alanine-D-alanyl carrier protein ligase (dcl) [imported]  
 C/Species: *Listeria innocua*  
 C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C/Accession: AD1554  
 R/Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
 D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A/Authors: Kretz, U.; Kuhn, M.; Kunst, F.; Kurapkak, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schlueker, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.  
 A/Title: Comparative genomics of *Listeria* species.  
 A/Reference number: AB1077; MUID:21537279; PMID:11679669  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-510 <G>  
 A/Cross-references: UNIPROT:Q92D47; GB:AL592022; PIDN:CAC96204.1; PID:G16413332; GSPDB: A/Experimental source: strain Clp11262  
 C/Genetics:  
 A/Gene: dclA

Query Match 73.8%; Score 31; DB 2; Length 510;  
 Best Local Similarity 55.6%; Pred. No. 78;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
 Db 224 TITANMKDL 232

## RESULT 6

transport protein homolog lln0896 [imported] - *Listeria innocua* (strain Clp11262)  
 C/Species: *Listeria innocua*  
 C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C/Accession: AF1544  
 R/Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
 D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A/Authors: Kretz, U.; Kuhn, M.; Kunst, F.; Kurapkak, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schlueker, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.  
 A/Title: Comparative genomics of *Listeria* species.  
 A/Reference number: AB1077; MUID:21537279; PMID:11679669  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-544 <G>  
 A/Cross-references: UNIPROT:Q92DB9; GB:AL592022; PIDN:CAC96128.1; PID:G16413346; GSPDB: A/Experimental source: strain Clp11262  
 C/Genetics:  
 A/Gene: lln0896  
 C/Superfamily: integral membrane protein HP0228

Query Match 73.8%; Score 31; DB 2; Length 544;  
 Best Local Similarity 66.7%; Pred. No. 84;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
 Db 154 SFVANKKEI 162

## RESULT 7

chaperonin 60 alpha chain precursor, chloroplast - rae  
 C/Species: *Brassica napus* (rape)  
 C/Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004  
 C/Accession: S38642  
 R/Cole, K.P.; Blakeley, S.D.; Dennis, D.T.  
 submitted to the EMBL Data Library, November 1993  
 A/Description: Isolation of a full length cDNA encoding *Brassica napus* plastid chaperon

A:Reference number: S38636  
 A:Accession: S38642  
 A:Molecule type: mRNA  
 A:Residues: 1-583 <COL>  
 A:Cross-references: UNIPROT:P34794; EMBL:Z27222; NID:G415924; PIDD:CA61736.1; PIDD:G4159  
 C:Genetics:  
 A:Genome: nuclear  
 C:Superfamily: chaperonin groEL  
 C:Keywords: chaperonin; heat shock; molecular chaperone  
 F:1-45/Domain: transit peptide (chloroplast) #status predicted <TMP>  
 F:45-583/Product: chaperonin 60 alpha chain, chloroplast #status predicted <MAT>

Query Match 73.8%; Score 31; DB 2; Length 583;  
 Best Local Similarity 66.7%; Pred. No. 91;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LTANKEV 9  
 DB 42 SLTANKEV 50

RESULT 9  
 F82291  
 conserved hypothetical protein VC0689 [imported] - Vibrio cholerae (strain N16961 serogr  
 C:Species: Vibrio cholerae  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
 C:Accession: F82291  
 R:Heideberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.U.;  
 Charatson, D.; Ermolova, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.  
 1, R.R.; Mekalanos, J.D.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
 A:Reference number: A82035; MUID:20406833; PMID:10952301  
 A:Accession: F82291  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-599 <HEI>  
 A:Cross-references: UNIPROT:Q9KU40; GB:AE004155; GB:AE003852; NID:G9655127; PIDD:AA9385  
 A:Experimental source: serogroup O1, strain N16961; biotype El Tor  
 C:Genetics:  
 A:Gene: VC0689  
 A:Map position: 1

Query Match 73.8%; Score 31; DB 2; Length 599;  
 Best Local Similarity 66.7%; Pred. No. 93;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LTANKEV 9  
 DB 76 SLTANKEV 84

RESULT 9  
 AF1259  
 transcription regulator (Merr family) homolog lmo1478 [imported] - Listeria monocytogene  
 C:Species: Listeria monocytogenes  
 C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C:Accession: AF1259  
 R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
 ; Dominguez-Serna, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fehl, H.  
 D.; Jones, L.M.; Karet, U.  
 Science 294, 849-852, 2001  
 A:Authors: Kretz, J.; Kuhn, M.; Kuntz, F.; Kurapat, G.; Madueno, E.; Matcounan, A.; Ma  
 ck, C.; Schluter, T.; Simoes, N.; Tiller, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund,  
 A:Title: Comparative genomics of Listeria species.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AF1259  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-145 <GLA>  
 A:Cross-references: UNIPROT:Q8Y743; GB:NC\_003210; PIDD:CA09556.1; PIDD:G16410907; GSPDB:  
 A:Experimental source: strain EGD-e  
 C:Genetics:

A:Gene: lmo1478

Query Match 71.4%; Score 30; DB 2; Length 145;  
 Best Local Similarity 75.0%; Pred. No. 32;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LTANKEV 9  
 DB 100 LTANKEV 107

RESULT 10  
 C95196  
 conserved hypothetical protein SPI685 [imported] - Streptococcus pneumoniae (strain TIG  
 C:Species: Streptococcus pneumoniae  
 C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004  
 C:Accession: C95196  
 R:Petelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Hei  
 on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,  
 nson, T.; Hickey, E.K.; Holt, I.E.  
 Science 293, 498-506, 2001  
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
 A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
 A:Reference number: A95000; MUID:21357209; PMID:11463916  
 A:Accession: C95196  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-232 <KUR>  
 A:Cross-references: UNIPROT:Q8DNU6; GB:AE005672; PIDD:AAK75764.1; PIDD:G14973178; GSPDB:  
 A:Experimental source: strain TIGR4  
 C:Genetics:  
 A:Gene: SPI685

Query Match 71.4%; Score 30; DB 2; Length 232;  
 Best Local Similarity 75.0%; Pred. No. 54;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LTANKEV 9  
 DB 86 ITATKEV 93

RESULT 11  
 H98062  
 N-acetylmannosamine-6-P epimerase [imported] - Streptococcus pneumoniae (strain R6)  
 C:Species: Streptococcus pneumoniae  
 C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
 C:Accession: H98062  
 R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.;  
 e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.D.; Lu, J.; Matsushima, P.; McAhren, S.;  
 Y, P.; Sun, P.M.; Winkler, M.E.  
 J. Bacteriol. 183, 5709-5717, 2001  
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Taskunas, S.R.  
 A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
 A:Reference number: A97872; MUID:21429245; PMID:11544234  
 A:Accession: H98062  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-232 <KUR>  
 A:Cross-references: UNIPROT:Q8DNU6; GB:AE007317; PIDD:AAU0033.1; PIDD:G15459192; GSPDB:  
 A:Genetics:  
 A:Gene: nanE

Query Match 71.4%; Score 30; DB 2; Length 232;  
 Best Local Similarity 75.0%; Pred. No. 54;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LTANKEV 9  
 DB 86 ITATKEV 93

RESULT 12

C98021  
N-acetylmannosamine-6-P epimerase [imported] - Streptococcus pneumoniae (strain R6)  
C/Species: Streptococcus pneumoniae  
C/Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C/Accession: C98021  
R/Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.C.; Lu, J.; Matsushima, P.; McAhren, S.; M  
y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A/Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;  
A/Title: Genome of the Bacterium Streptococcus pneumoniae strain R6.  
A/Reference number: A97872; MUID:21429245; PMID:11544234  
A/Accession: C98021  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-232 <KUR>  
A/Cross-references: UNIPROT:Q8DPFO; GB:AE007317; PIDD:AAK99999.1; PID:G15458829; GSPDB:G  
C/Genetics:  
A/gene: nanB

Query Match 71.4%; Score 30; DB 2; Length 232;  
Best Local Similarity 75.0%; Pred. No. 54;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LTANKKEV 9  
||| |||  
Db 86 ITANKKEV 93

RESULT 13  
C95154  
N-acetylmannosamine-6-P epimerase, probable [imported] - Streptococcus pneumoniae (strain  
C/Species: Streptococcus pneumoniae  
C/Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004  
C/Accession: C95154  
R/Hetzel, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Kadane, D.; Holtapple,  
neon, T.; Hickey, E.K.; Holt, I.B.  
Science 293, 496-506, 2001  
A/Authors: Loftus, B.J.; Yang, F.; Smith, H.C.; Venter, J.C.; Dougherty, B.A.; Morrison,  
A/Title: Complete genome sequence of a virulent isolate of Streptococcus pneumoniae.  
A/Reference number: A95000; MUID:21357209; PMID:11463916  
A/Accession: C95154  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-233 <KUR>  
A/Cross-references: UNIPROT:Q97095; GB:AE005672; PIDD:AAK75428.1; PID:G14972812; GSPDB:G  
A/Experimental source: strain ITGR4  
C/Genetics:  
A/gene: SPI330

Query Match 71.4%; Score 30; DB 2; Length 233;  
Best Local Similarity 75.0%; Pred. No. 54;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LTANKKEV 9  
||| |||  
Db 87 ITANKKEV 94

RESULT 14  
A86732  
geranyltransferase (EC 2.5.1.10) [imported] - Lactococcus lactis subsp. lactis (str  
N/Alternative names: farnesyl diphosphate synthase  
C/Species: Lactococcus lactis subsp. lactis  
C/Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 09-Jul-2004  
C/Accession: A86732  
R/Bolotin, A.; Winkler, P.; Mauger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrli  
Genome Res. 11, 731-753, 2001  
A/Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s  
A/Reference number: A86625; MUID:21235186; PMID:11337471  
A/Accession: A86732  
A/Status: preliminary

A/Molecule type: DNA  
A/Residues: 1-285 <STO>  
A/Cross-references: UNIPROT:Q9CH81; GB:AE005176; PIDD:G12723785; PIDD:AAK04955.1; GSPDB:G  
A/Experimental source: strain ILL403  
C/Genetics:  
A/gene: ispA  
C/Keywords: transferase

Query Match 71.4%; Score 30; DB 2; Length 285;  
Best Local Similarity 66.7%; Pred. No. 68;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 SLTANKKEV 9  
||| |||  
Db 252 SLTANKKEV 260

RESULT 15  
S61498  
chemotactic methyltransferase homolog pilK - Pseudomonas aeruginosa  
C/Species: Pseudomonas aeruginosa  
C/Date: 27-Apr-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C/Accession: S61498; D83594  
R/Darzens, A.  
Mol. Microbiol. 15, 703-717, 1995  
A/Title: The Pseudomonas aeruginosa pilK gene encodes a chemotactic methyltransferase ('  
A/Reference number: S61498; MUID:95302983; PMID:7783642  
A/Accession: S61498  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-291 <DAR>  
A/Cross-references: UNIPROT:Q51346; EMBL:U11382; NID:G520528; PIDD:AA85790.1; PID:G5205  
R/Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Laidig, K.; Lim,  
Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path  
A/Reference number: A82950; MUID:20437337; PMID:10964043  
A/Accession: D83594  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-291 <STO>  
A/Cross-references: GB:AE004478; GB:AE004091; NID:G9946261; PIDD:AA803801.1; GSPDB:GN001  
A/Experimental source: strain PA01  
C/Genetics:  
A/gene: pilK; PA0412  
A/superfamily: protein-glutamate O-methyltransferase; protein-glutamate O-methyltransfer  
F;3-274/Domain: protein-glutamate O-methyltransferase homology <PKX>

Query Match 71.4%; Score 30; DB 2; Length 291;  
Best Local Similarity 66.7%; Pred. No. 70;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SLTANKKEV 9  
||| |||  
Db 46 SLTANKKEV 54

Search completed: November 30, 2004, 07:41:26  
Job time: 20 secs

Tue Nov 30 08:49:33 2004

us-10-008-377a-7.open.rup

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 : Search time 68.7692 Seconds  
(without alignments)  
75.301 Million cell updates/sec

Title: US-10-008-377a-7  
Perfect score: 42  
Sequence: 1 SLTANKKEV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 segs, 575374646 residues  
Total number of hits satisfying chosen parameters: 1825181

Minimum DB seg length: 0  
Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_02:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	100.0	803	2 Q8TDJ5	Q8TDJ5 homo sapien
2	42	100.0	1620	1 ALK_HUMAN	Q9UM73 homo sapien
3	35	83.3	1194	2 Q9R6P0	Q9R6P0 agrobacteri
4	34	81.0	286	2 Q8A0P1	Q8A0P1 bacteroides
5	33	78.6	206	2 Q20521	Q20521 caenorhabdi
6	33	78.6	660	2 Q17248	Q17248 boophilus m
7	33	78.6	807	2 Q81B74	Q81B74 plasmodium
8	32	76.2	320	2 Q6TV42	Q6TV42 bacillus me
9	32	76.2	320	2 AAR39404	AAR39404 bacillus
10	32	76.2	510	2 Q8W2N4	Q8W2N4 nistecia t
11	32	76.2	541	2 Q8Y8K4	Q8Y8K4 nistecia mo
12	32	76.2	684	2 Q6ZK85	Q6ZK85 cryza sativ
13	32	76.2	684	2 BAD08825	BAD08825 cryza sat
14	31	73.8	124	2 Q74CFE	Q74CFE geobacter s
15	31	73.8	124	2 AAR35086	AAR35086 geobacter
16	31	73.8	170	2 Q6K262	Q6K262 picophilus
17	31	73.8	238	2 Q6T2V4	Q6T2V4 homo sapien
18	31	73.8	238	2 AAR13901	AAR13901 homo sapi
19	31	73.8	250	2 Q8N4X4	Q8N4X4 homo sapien
20	31	73.8	254	2 Q8BGU6	Q8BGU6 m mus muscu
21	31	73.8	309	2 Q8GY34	Q8GY34 arabidopsis
22	31	73.8	367	2 Q839S1	Q839S1 enterococcu
23	31	73.8	510	1 DLRN_LISTN	DLRN_LISTN nistecia in
24	31	73.8	510	1 DLRN_LISTN	DLRN_LISTN nistecia in
25	31	73.8	510	1 Q72HJ2	Q72HJ2 nistecia mo
26	31	73.8	510	2 AAT03772	AAT03772 nistecia
27	31	73.8	544	2 Q92DB9	Q92DB9 nistecia in
28	31	73.8	583	1 RUPA_BRANA	RUPA_BRANA brassica na
29	31	73.8	583	1 Q6BVC7	Q6BVC7 debaryomyce
30	31	73.8	598	2 Q73HR8	Q73HR8 wolbachia p
31	31	73.8	598	2 AAS14195	AAS14195 wolbachia

32	31	73.8	599	2 Q87JH6	Q87JH6 vibrio para
33	31	73.8	599	2 Q9KU40	Q9KU40 vibrio chol
34	31	73.8	764	2 Q76130	Q76130 neuropeptid
35	31	73.8	862	1 DR33_ARATH	DR33_ARATH arabidopsis
36	31	73.8	1113	2 Q6CMFO	Q6CMFO kiuyveromyc
37	30	71.4	110	2 Q6EHL3	Q6EHL3 biophthalari
38	30	71.4	133	2 Q9N3J4	Q9N3J4 caenorhabdi
39	30	71.4	145	2 Q8Y743	Q8Y743 nistecia mo
40	30	71.4	145	2 Q71Z02	Q71Z02 nistecia mo
41	30	71.4	149	2 AAT04272	AAT04272 nistecia
42	30	71.4	179	1 ATPD_STAEP	ATPD_STAEP streptococ
43	30	71.4	212	2 Q6NN76	Q6NN76 dirosophila
44	30	71.4	212	2 AAR62209	AAR62209 dirosophila
45	30	71.4	232	1 NAEI_STR6	Q8P60 streptococc

#### ALIGNMENTS

```
RESULT 1
ID Q8TDJ5 PRELIMINARY; PRI; 803 AA.
AC Q8TDJ5;
DT 01-JUN-2002 (TREMBL) 21, Created
DT 01-JUN-2002 (TREMBL) 21, Last sequence update
DT 01-MAR-2004 (TREMBL) 26, Last annotation update
DE TRK-fused gene/anaplastic large cell lymphoma kinase extra long
DE form.
GN Name=TRK/ALK fusion;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21940189; PubMed=11943732;
RA Hernandez L., Bea S., Bellosillo B., Pinyol M., Falini B., Ott G.,
RA Pulford K., Rosenwald A., Morris S.W., Fernandez A., Santos E.,
RA Campo E.;
RT "Diversity of genomic breakpoints in TRK-ALK translocations in
RT anaplastic large cell lymphomas: identification of a new TRK-ALK (XL)
RT chimeric gene with transforming activity."
RL Am. J. Pathol. 160:1487-1494 (2002).
DR EMBL: AF390893; AAM17922.1; -
DR HSP: Q62839; ILUP.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0004723; F:Protein-tyrosine kinase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR011009; Kinase-like.
DR InterPro: IPR000270; OPR_PBI.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002011; ReceptTyknsit.
DR InterPro: IPR001245; Tyr_kinase.
DR InterPro: IPR008266; Tyr_kinase_AS.
DR Pfam: PF00564; PBI.1.
DR Pfam: PF00069; Kinase.1.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: PD000001; Prot_kinase; 1.
DR SMART: SMO0666; PBI.1.
DR SMART: SMO0219; TYKc.1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00239; RECEPTOR_TYR_KIN_TYR; UNKNOWN_1.
DR ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.
SQ
SEQUENCE 803 AA; 88670 MW; 8908264356595F CRC64;
Query Match 100.0%; Score 42; DB 2; Length 803;
Best local similarity 100.0%; Pred. No. 2.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

DB 744 SLTANKEV 752

```

RESULT 2
ID ALK_HUMAN STANDARD; PRT; 1620 AA.
AC Q9UM73; Q9Y4K6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 05-UL-2004 (Rel. 44, Last annotation update)
DE ALK tyrosine kinase receptor precursor (BC 2.7.1.112) (Anaplastic
DE lymphoma kinase) (CD246 antigen).
DE Name:ALK;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, TISSUE SPECIFICITY, AND
RP GLYCOSYLATION.
RX MEDLINE=97316779; PubMed=9174053; DOI=10.1038/sj.onc.1201062;
RA Morris S.W., Naeye C.W., Mathew P., James P.L., Kirstein M.N., Cui X.,
RA Witte D.P.;
RT "ALK, the chromosome 2 gene locus altered by the t(2;5) in non-
RT Hodgkin's lymphoma, encodes a novel neural receptor tyrosine kinase
RT that is highly related to leukocyte tyrosine kinase (LTK).";
RL Oncogene 14:2175-2188(1997).
RN [2]
RP ERRATUM.
RA Morris S.W., Naeye C.W., Mathew P., James P.L., Kirstein M.N., Cui X.,
RA Witte D.P.;
RL Oncogene 15:2883-2883(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97178863; PubMed=9053841;
RA Iwabara T., Fujimoto J., Ken D.,
RA Mori S., Katzkin B., Yamamoto T.;
RT "Molecular characterization of ALK, a receptor tyrosine kinase
RT expressed specifically in the nervous system.";
RL Oncogene 14:439-449(1997).
RN [4]
RP PARTIAL SEQUENCE FROM N.A., AND CHROMOSOMAL TRANSLOCATION.
RX MEDLINE=94167588; PubMed=8122112;
RA Morris S.W., Kirstein M.N., Valentine M.B., Dittmer K.G.,
RA Shapiro D.N., Saltman D.L., Look A.T.;
RT "Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in
RT non-Hodgkin's lymphoma.";
RL Science 263:1281-1284(1994).
CC -1- FUNCTION: Orphan receptor with a tyrosine-protein kinase activity.
CC Appears to play an important role in the normal development and
CC function of the nervous system.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed in brain and CNS. Also expressed in
CC the small intestine and testis, but not in normal lymphoid cells.
CC -1- PTM: N-glycosylated.
CC -1- DISEASE: A form of non-Hodgkin's lymphoma is characterized by a
CC chromosomal translocation t(2;5)(p23;q35) that involves NPM1 and
CC ALK.
CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin
CC receptor subfamily.
CC -1- SIMILARITY: Contains 1 LDM-receptor class A domain.
CC -1- SIMILARITY: Contains 2 WML domains.
CC -1- DATABASE: NMRAtlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW="http://www.inbioigen.fr/services/cytocancer/Genes/ALK.html".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC EMBL; U62540; AAB71619.1; -
DR EMBL; U66559; AAC51104.1; -
DR HSSP; Q62838; IJUF.
DR Genew: HGNC:427; ALK.
DR MIM; 105590; .
DR GO; GO:0016021; C: integral to membrane; NAS.
DR GO; GO:007359; P: neurogenesis; NAS.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000988; MAM.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002011; Recepttyrkinst.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF00629; MAM; 1.
DR Pfam; PF00063; Kinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00192; LDla; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS01209; LDLRA_1; FALSE NEG.
DR PROSITE; PS00068; LDLRA_2; FALSE NEG.
DR PROSITE; PS00740; MAM_1; FALSE NEG.
DR PROSITE; PS00060; MAM_2; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
DR ATP-binding; Chromosomal translocation; Glycoprotein; Phosphorylation;
KW Proto-oncogene; Receptor; Repeat; Signal; Transferase; Transmembrane;
KW Tyrosine-protein kinase.
FT SIGNAL 1 18
FT CHAIN 19 1620
FT DOMAIN 19 1038
FT TRANSXSM 1039 1055
FT DOMAIN 1060 1620
FT DOMAIN 264 427
FT DOMAIN 437 473
FT DOMAIN 478 636
FT DOMAIN 1116 1392
FT DOMAIN 816 940
FT NP_BIND 1122 1130
FT BINDING 1150 1150
FT ACT_SITE 1249 1249
FT MOD_RES 1282 1282
FT CARBOHYD 169 169
FT CARBOHYD 244 244
FT CARBOHYD 285 285
FT CARBOHYD 324 324
FT CARBOHYD 411 411
FT CARBOHYD 424 424
FT CARBOHYD 445 445
FT CARBOHYD 563 563
FT CARBOHYD 571 571
FT CARBOHYD 627 627
FT CARBOHYD 709 709
FT CARBOHYD 808 808
FT CARBOHYD 863 863
FT CARBOHYD 864 864
FT CARBOHYD 886 886
FT CARBOHYD 986 986
FT CONFLICT 36 36
FT CONFLICT 1491 1491
FT CONFLICT 1529 1529
SQ SEQUENCE 1620 AA; 176417 MW; A626049824296151E CRC64;

```

Query Match 100.0%; Score 42; DB 1; Length 1620;  
 Best Local Similarity 100.0%; Pred. No. 6.1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
 DB 1561 SLTANKEV 1569

## RESULT 3

Q9R6F0 PRELIMINARY; PRT: 1194 AA.  
 AC Q9R6F0;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Tior109 protein.  
 GN Name=Tior109;  
 OS Agrobacterium tumefaciens.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 CC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
 CX NCBI\_TaxID=558;

SEQUENCE FROM N.A.  
 RC STRAIN=MAF301001;  
 RA Hattori Y., Suzuki K., Ohta N., Uraji M., Katoh A., Yoshida K.;  
 RT "Genome structure of pTI-SAKURA (II): Strategy for DNA sequencing of a  
 RT Japanese cherry-Ti plasmid.";  
 RL Nucleic Acids Symp. Ser. 37:159-160 (1998).

SEQUENCE FROM N.A.  
 RC STRAIN=MAF301001;  
 RA Ohta N., Suzuki K., Hattori Y., Uraji M., Katoh A., Yoshida K.;  
 RT "Genome structure of pTI-SAKURA (III): Characteristics of T-DNA.";  
 RL Nucleic Acids Symp. Ser. 39:185-186 (1998).

SEQUENCE FROM N.A.  
 RC STRAIN=MAF301001;  
 RA Uraji M., Suzuki K., Ohta N., Hattori Y., Katoh A., Yoshida K.;  
 RT "Genome structure of pTI-SAKURA (IV): Characteristics of tra region.";  
 RL Nucleic Acids Symp. Ser. 39:187-188 (1998).

SEQUENCE FROM N.A.  
 RC STRAIN=MAF301001;  
 RA Hattori Y., Suzuki K., Ohta N., Uraji M., Katoh A., Yoshida K.;  
 RT "Genome structure of pTI-SAKURA (V): Complete nucleotide sequence of  
 RT plasmid pTI-SAKURA's vir region in Agrobacterium tumefaciens.";  
 RL Nucleic Acids Symp. Ser. 39:265-266 (1998).

SEQUENCE FROM N.A.  
 RC STRAIN=MAF301001;  
 RX MEDLINE=20184752; PubMed=10721272;  
 RA Suzuki K., Hattori Y., Uraji M., Ohta N., Iwata K., Murata K.,  
 RA Katoh A., Yoshida K.;  
 RT "Complete nucleotide sequence of a plant tumor-inducing Ti plasmid.";  
 RL Gene 242:331-336 (2000).

SEQUENCE FROM N.A.  
 RC STRAIN=MAF301001;  
 RX MEDLINE=9813120; PubMed=9524202;  
 RA Suzuki K., Ohta N., Hattori Y., Uraji M., Katoh A., Yoshida K.;  
 RT "Novel structural difference between nopaline- and octopine- type trbT  
 RT gene: construction of genetic and physical map and sequencing of  
 RT trbT/traI and rep gene clusters of a new Ti plasmid pTI-SAKURA.";  
 RL Biochim. Biophys. Acta 1396:1-7 (1998).

EMBL: AB016260; BAA8734.1; -;  
 GO: GO:0009291; Pseudocirculn conjugation; IEA.

InterPro: IPR005053; ModA\_ModL.  
 Pfam: Pf03389; ModA\_ModL; 1.

Query Match 83.3%; Score 35; DB 2; Length 1194;  
 Best Local Similarity 77.8%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
 DB 920 SLTANKEV 928

## RESULT 4

Q9A0P1 PRELIMINARY; PRT: 256 AA.  
 AC Q9A0P1;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocustNames=BT3980;  
 OS Bacteroides thetaiotaomicron.  
 OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;  
 CC Bacteroidaceae; Bacteroides.  
 CX NCBI\_TaxID=818;

SEQUENCE FROM N.A.  
 RC STRAIN=VPI-5482 / ATCC 29148;  
 RX MEDLINE=22550858; PubMed=1263928;  
 RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,  
 RA Chang H.C., Hooper L.V., Gordon J.I.;  
 RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";  
 RL Science 289:2074-2076 (2003).

EMBL: AB016943; AA079085.1; -;  
 GO: GO:0006943; Hypothetical protein.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 256 AA; 30314 MW; 6EF4EA7649790EDC CRC64;

Query Match 81.0%; Score 34; DB 2; Length 256;  
 Best Local Similarity 77.8%; Pred. No. 49;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
 DB 46 SLTANKEV 54

## RESULT 5

Q20521 PRELIMINARY; PRT: 206 AA.  
 AC Q20521;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein F47B8.8.  
 GN ORFNames=F47B8.8;  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 CC Rhabditidae; Petodermidae; Caenorhabditis.  
 CX NCBI\_TaxID=6239;

SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology.";  
 RL Science 282:2012-2018 (1998).

SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Beiks M.;  
 RT Sumitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: Z77662; CAB01196.1; -;  
 DR PIR: T22345; T22345.

WormPep: F47B8.8; CE10646.  
 DR GO: GO:004179; F-membrane alanyl aminopeptidase activity; IEA.  
 DR GO: GO:0006508; Proteolysis and peptidolysis; IEA.  
 DR InterPro: IPR001930; Peptidase M1.  
 Pfam: Pf01433; Peptidase\_M1; 1.

KW Hypothetical protein.  
SQ SEQUENCE 206 AA; 23204 MW; A1952AA79BD86075 CRC64;

Query Match 78.6%; Score 33; DB 2; Length 206;  
Best Local Similarity 75.0%; Pred. No. 65;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKE 8  
DB 131 SLTANKE 138

## RESULT 6

Q17248 PRELIMINARY; PRT; 660 AA.  
AC 017248;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)  
DE Angiotensin-converting enzyme-like protein precursor.  
GN Name=Ang2;  
OS Boophilus microplus (Cattle tick).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Parasitiformes; Ixodidae; Ixodidae; Boophilus.  
OX NCBI\_TaxID=6941;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-whole ticks;  
RA Whitefield P.L., Johnson M.C., Smith D.R.J., Snelson V.J., Clifton G.F.,  
RA Brown G.S., Cairns D., Foy A.B., Irving D.O.;  
RT Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U62809; AAB04998.1; -  
DR HSSP; Q10714; IJ36.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0004246; F:peptidyl-dipeptidase A activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR001546; Peptidase\_M2.  
DR Pfam; PF01401; Peptidase\_M2; 1.  
DR PRINTS; PRO0791; PEPDPTASEA.  
KW Signal.  
FT CHAIN 1 29 Potential.  
FT SIGNAL 30 660 Potential.  
SQ SEQUENCE 660 AA; 75257 MW; 6F164CF70C938E63 CRC64;

Query Match 78.6%; Score 33; DB 2; Length 660;  
Best Local Similarity 87.5%; Pred. No. 2,4e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LTANKEV 9  
DB 173 LTANKEV 180

## RESULT 7

Q81B74 PRELIMINARY; PRT; 807 AA.  
AC 081B74;  
DT 01-MAR-2003 (TREMBlrel. 23, Created)  
DT 01-MAR-2003 (TREMBlrel. 23, last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, last annotation update)  
DE Hypothetical protein PF08\_0030.  
GN Name=PF08\_0030;  
OS Plasmodium falciparum (isolate 3D7).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.  
OX NCBI\_TaxID=36629;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC Seegeer K., Murphy L., Harris D., Berrihan M., Pain A., Hall N.,  
RA Quail M., Barrett B.;  
RT Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL844507; CAD51133.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 807 AA; 95820 MW; 7CFB4D3A720625F2 CRC64;

Query Match 76.6%; Score 33; DB 2; Length 807;  
Best Local Similarity 87.5%; Pred. No. 2.9e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLTANKE 8  
DB 764 SLTANKE 771

## RESULT 8

Q6TV42 PRELIMINARY; PRT; 320 AA.  
AC 06TV42;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, last annotation update)  
DE Putative class II fructose-1,6-bisphosphatase / seduhexulose-1,7-bisphosphatase.  
GN Name=gipx;  
OS Bacillus methanolicus.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1471;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MG3;  
RC PubMed:14973041;  
RA Brautaset T., Jakobsen O.M., Flickinger M.C., Valla S.,  
RA Ellingsen T.E.;  
RT "Plasmid-dependent methylotrophy in thermotolerant Bacillus methanolicus."  
RT J. Bacteriol. 186:1229-1238(2004).  
DR EMBL; AY386314; AAR39404.1; -  
DR InterPro; IPR002453; Beta\_tubulin.  
DR InterPro; IPR004454; GIPX.  
DR Pfam; PF003320; FBpase\_gipx; 1.  
DR Pfam; PF007014; GIPX; 1.  
DR TIGRFAMs; TIGR00330; GIPX; 1.  
DR PROSITE; PS00228; TUBULIN\_B\_AUTOREG; UNKNOWN\_1.  
KW Plasmid.  
SQ SEQUENCE 320 AA; 34557 MW; 816B36237154628C CRC64;

Query Match 76.2%; Score 32; DB 2; Length 320;  
Best Local Similarity 77.8%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
DB 146 SLTANKEV 154

## RESULT 9

AAR39404 PRELIMINARY; PRT; 320 AA.  
AC AAR39404;  
DT 02-MAR-2004 (TREMBlrel. 27, Created)  
DT 02-MAR-2004 (TREMBlrel. 27, last sequence update)  
DT 02-MAR-2004 (TREMBlrel. 27, last annotation update)  
DE Putative class II fructose-1,6-bisphosphatase / seduhexulose-1,7-bisphosphatase.  
GN GIPX.  
OS Bacillus methanolicus.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1471;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MG3;  
RC PubMed:14973041;  
RA Brautaset T., Jakobsen O.M., Flickinger M.C., Valla S.,  
RA Ellingsen T.E.;  
RT "Plasmid-dependent Methylotrophy in Thermotolerant Bacillus

```

RT methanolicus."
RL J. Bacteriol. 186:1229-1238(2004).
DR EMBL: AY386314; MAR39404.1; -.
KW plasmid.
SQ SEQUENCE 320 AA; 34557 MW; 81EE36237154628C CRC64;

Query Match
Best Local Similarity 76.2%; Score 32; DB 2; Length 320;
Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTANKEV 9
DB 146 SLTANKEV 154

RESULT 10
Q8W2N4 PRELIMINARY; PRT; 510 AA.
AC Q8W2N4;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Cytochrome P450-dependent fatty acid hydroxylase.
GN Name=CYP9A6;
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RA Lehouquin R., Kahn R., Benveniste I., Durst F.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL: AF092913; AAL54884.1; -.
DR GO: GO:0004497; F:monooxygenase activity; IEA.
DR GO: GO:0006118; P:electron transport; IEA.
DR InterPro: IPR001128; Cytochrome_P450.
DR InterPro: IPR002401; EP4501.
DR Pfam: PF00067; P450; 1.
DR PRINTS: PR00463; EP4501.
DR PRINTS: PR00385; P450.
KM Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 510 AA; 58633 MW; 67F0FD6472DA182 CRC64;

Query Match
Best Local Similarity 76.2%; Score 32; DB 2; Length 510;
Pred. No. 3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LITANKEV 9
DB 494 LITANKEV 501

RESULT 11
Q8Y8K4 PRELIMINARY; PRT; 541 AA.
AC Q8Y8K4;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Lmo0897 protein.
GN OrderedLocName=Lmo0897;
OS Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=1679669;
RA Glaser P., Francau L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bioecker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Darvar A., Dehoux P.,

```

```

RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fehli H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hant U., Jackson D.,
RA Jones L.-M., Kaerst U., Krell T., Kuhn M., Kunst F., Kurapkak G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RT "Comparative genomics of Listeria species."
RL Science 294:849-852(2001).
DR EMBL: AL581977; CAC8975.1; -.
DR FIR: A1186; A1186.
DR ListLib: LMO0897; -.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0008271; P:sulfate porter activity; IEA.
DR GO: GO:0008272; P:sulfate transport; IEA.
DR InterPro: IPR002845; SPAS.
DR InterPro: IPR01547; Sulph_transp.
DR InterPro: IPR001902; SulP_transp.
DR Pfam: PF01740; STAS; 1.
DR Pfam: PF00916; Sulfate_transp; 1.
DR TIGRfam: TIGRfam0815; sulP; 1.
DR PROSITE: PS50801; STAS; 1.
KM Complete proteome.
SQ SEQUENCE 541 AA; 58261 MW; B3A3A1A76A36EF2A CRC64;

Query Match
Best Local Similarity 76.2%; Score 32; DB 2; Length 541;
Pred. No. 3.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTANKEV 9
DB 154 SEVANKEV 162

RESULT 12
Q6ZK85 PRELIMINARY; PRT; 684 AA.
AC Q6ZK85;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DR RNA recognition motif (RRM)-containing protein-like.
GN Name=CJ1134 H03.9-1;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhacoidae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP003983; BAB08825.1; -.
DR InterPro: IPR005054; RNA_rec_mot.
DR Pfam: PF00076; RRM; 1.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PS50102; RRM; 1.
SQ SEQUENCE 684 AA; 75299 MW; 0E3C3F5DE072FDF8 CRC64;

Query Match
Best Local Similarity 76.2%; Score 32; DB 2; Length 684;
Pred. No. 4.1e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9
DB 486 SLTANKEV 494

RESULT 13
BAB08825 PRELIMINARY; PRT; 684 AA.
AC BAB08825;

```

DT 02-MAR-2004 (TREMBlrel. 27, Created)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)  
 DE RNA recognition motif (RNM)-containing protein-like.  
 GN OJ134\_H03-9-1.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eriatridae; Oryzaceae; Oryza; Oryza sativa.  
 OC NCBI\_TaxID=39947;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GAS) genomic DNA, chromosome 8, BAC  
 clone:OJ134\_H03-9-1";  
 RU Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, AP003883, BAD08825.1, OE3C3F5DE072F08F CRC64;  
 SQ SEQUENCE 684 AA; 75299 MW; 0E3C3F5DE072F08F CRC64;

Query Match 76.2%; Score 32; DB 2; Length 684;  
 Best Local Similarity 66.7%; Pred. No. 4.1e+02;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 LTANKEV 9  
 DB 486 SLISNIKL 494

RESULT 14  
 ID Q7AF5 PRELIMINARY; PRT; 124 AA.  
 AC Q7AF5;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Hypothetical protein.  
 GN ORFNames=GSU1719;  
 OS Geobacter sulfurreducens.  
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;  
 OC Geobacteraceae; Geobacter.  
 OC NCBI\_TaxID=35554;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PCA / ATCC 51573;  
 RA Pubmed=14671304; DOI=10.1126/science.1086727;  
 RA Methe B.A., Nelson K.E., Eisen J.A., Paulsen I.T., Nelson W.C.,  
 RA Heidelberg J.F., Wu D., Ward N.L., Beanan M.J., Dodson R.J.,  
 RA Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,  
 RA Gwinn M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,  
 RA Daviden T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A.,  
 RA Weidman J., Khoult H.M., Feldlyum T.V., Utterback T.R.,  
 RA Van Aken S.E., Lovley D.R., Fraser C.M.;  
 RT "Genome of Geobacter sulfurreducens: metal reduction in subsurface  
 RT environments.";  
 RL Science 302:1967-1969(2003).  
 DR EMBL, AE017213; AAR35096.1; -.  
 DR TIGR; GSU1719; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 124 AA; 13279 MW; 0098A8E5109A24DA CRC64;

Query Match 73.8%; Score 31; DB 2; Length 124;  
 Best Local Similarity 65.7%; Pred. No. 1e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 LTANKE 8  
 DB 88 LTANKE 94

RESULT 15  
 AAR35096 PRELIMINARY; PRT; 124 AA.  
 ID AAR35096

AC AAR35096;  
 DT 02-MAR-2004 (TREMBlrel. 27, Created)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)  
 DE Hypothetical protein.  
 GN GSU1719.  
 OS Geobacter sulfurreducens.  
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;  
 OC Geobacteraceae; Geobacter.  
 OC NCBI\_TaxID=35554;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PCA / ATCC 51573;  
 RA Pubmed=14671304;  
 RA Methe B.A., Nelson K.E., Eisen J.A., Paulsen I.T., Nelson W.C.,  
 RA Heidelberg J.F., Wu D., Ward N.L., Beanan M.J., Dodson R.J.,  
 RA Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,  
 RA Gwinn M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,  
 RA Daviden T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A.,  
 RA Weidman J., Khoult H.M., Feldlyum T.V., Utterback T.R.,  
 RA Van Aken S.E., Lovley D.R., Fraser C.M.;  
 RT "Genome of Geobacter sulfurreducens: metal reduction in subsurface  
 RT environments.";  
 RL Science 302:1967-1969(2003).  
 DR EMBL, AE017213; AAR35096.1; -.  
 DR TIGR; GSU1719; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 124 AA; 13279 MW; 0098A8E5109A24DA CRC64;

Query Match 73.8%; Score 31; DB 2; Length 124;  
 Best Local Similarity 65.7%; Pred. No. 1e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 LTANKE 8  
 DB 88 LTANKE 94

Search completed: November 30, 2004, 07:17:32  
 Job time : 70.7692 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 ; Search time 69.4615 Seconds  
(without alignments)  
46,480 Million cell updates/sec

Title: US-10-008-377a-7  
Perfect score: 42  
Sequence: 1 SLTANKEV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues  
Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A.Genseq\_23sep04: \*  
1: genseqpl1808: \*  
2: genseqpl1908: \*  
3: genseqpl2008: \*  
4: genseqpl2001a: \*  
5: genseqpl2002a: \*  
6: genseqpl2003a: \*  
7: genseqpl2003b: \*  
8: genseqpl2004a: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	9	AA022989	AA022989 Human p62
2	42	100.0	675	ABG20950	ABG20950 Novel hum
3	42	100.0	680	AA013858	AA013858 Human NPM
4	42	100.0	680	ABG95067	ABG95067 Human tta
5	42	100.0	680	ABU04347	ABU04347 Human exp
6	42	100.0	680	ABU04362	ABU04362 Human exp
7	42	100.0	680	ABU04358	ABU04358 Human exp
8	42	100.0	680	ABG74472	ABG74472 NPM/ALK f
9	42	100.0	1620	AA013857	AA013857 Human ALK
10	42	100.0	1620	AA013858	AA013858 Human ins
11	42	100.0	1620	ABG74471	ABG74471 Human ALK
12	33	78.6	560	AA013858	AA013858 Tick carb
13	33	76.2	510	AA013858	AA013858 Tobacco f
14	33	76.2	541	AB047568	AB047568 Listeria
15	33	76.2	541	ABU32622	ABU32622 Protein e
16	31	73.8	41	AA025358	AA025358 Human pro
17	31	73.8	121	AA045378	AA045378 Human sec
18	31	73.8	136	AA013859	AA013859 Human pol
19	31	73.8	250	ADP13932	ADP13932 Human end
20	31	73.8	265	ADP13932	ADP13932 Human end
21	31	73.8	277	ADP13940	ADP13940 Human end
22	31	73.8	319	ADP13938	ADP13938 Human end
23	31	73.8	334	ABG16486	ABG16486 Novel hum
24	31	73.8	339	ABU25245	ABU25245 Protein e
25	31	73.8	378	ADH89725	ADH89725 Enterococ

26	31	73.8	510	AB048924	AB048924 Listeria
27	31	73.8	510	ABU32635	ABU32635 Protein e
28	31	73.8	940	ADP13937	ADP13937 Human end
29	31	73.8	1250	ADP13935	ADP13935 Human end
30	31	73.8	1335	ADP13936	ADP13936 Human end
31	31	73.8	1683	ADP13937	ADP13937 Mouse kin
32	31	73.8	2247	ADP13933	ADP13933 Human end
33	31	73.8	2383	ADP13934	ADP13934 Human end
34	31	73.8	2545	AB08406	AB08406 Human NOV
35	30	71.4	122	AA045377	AA045377 Gene 38 h
36	30	71.4	145	AB048805	AB048805 Listeria
37	30	71.4	179	AA082357	AA082357 S. epider
38	30	71.4	179	ABU43169	ABU43169 Protein e
39	30	71.4	180	ABP39370	ABP39370 Staphyloc
40	30	71.4	190	AA040628	AA040628 Human ORF
41	30	71.4	190	ABP00040	ABP00040 Human ORF
42	30	71.4	232	ABU02158	ABU02158 S. pneumo
43	30	71.4	232	ADK46663	ADK46663 Streptoco
44	30	71.4	232	ADK47957	ADK47957 Streptoco
45	30	71.4	233	ABU01769	ABU01769 S. pneumo

## ALIGNMENTS

RESULT 1	AA022989	standard; peptide; 9 AA.
ID	AA022989	
AC	AA022989	
XX		
DT	17-SEP-2003	(first entry)
XX		
DE	Human p62-629 ALK-derived HLA-A*0201 restricted CTL epitope peptide.	
XX		
XX	HLA-A*0201 binding; anaplastic lymphoma kinase; ALK; CTL epitope; ALCL;	
KW	cytotoxic T-lymphocyte; receptor tyrosine; NPM/ALK fusion protein; human;	
KW	oncogenic; anaplastic large cell lymphoma; neuroblastoma; gene therapy;	
KW	t(2;5) translocation; neoplasia; cytostatic; immunotherapy; alkoma;	
XX	p62-629.	
OS	Homo sapiens.	
XX		
PN	WO2003042243-A2.	
XX		
PD	22-MAY-2003.	
XX		
PF	14-NOV-2002; 2002WO-EP012764.	
XX		
PR	15-NOV-2001; 2001US-00008377.	
XX		
PA	(NAST-) 1ST NAZ STUDIO E CURA DEI TUMORI.	
PA	(STUD-) ST JUDE CHILDREN'S RES HOSPITAL.	
XX		
PI	Gambacorti-Passerini C, Passoni L;	
XX		
DR	WPI; 2003-441791/41.	
XX		
PT	New HLA-A*0201-Binding Anaplastic Lymphoma Kinase (ALK) peptide,	
PT	useful for preparing a composition for treating ALK-positive lymphoma,	
PT	neuroblastoma or ALK-expressing neoplasia.	
XX		
PS	Claim 1; Page 7, 33pp; English.	
XX		
CC	The invention relates to a novel HLA-A*0201-binding anaplastic lymphoma	
CC	kinase (ALK)-derived peptide, of 9-10 amino acids, which may act as a	
CC	cytotoxic T-lymphocyte (CTL)-epitope. ALK is a receptor tyrosine kinase	
CC	which forms part of the NPM/ALK fusion protein (oncogenic lymphoma kinase	
CC	fusion protein) in ALCL (anaplastic large cell lymphoma)-derived cell	
CC	lines. More than 50% of ALCL cases possess a t(2;5) chromosomal	
CC	translocation that leads to the expression of the NPM/ALK fusion protein	
CC	which forms a potent oncogene when constitutively activated. Translocated	
CC	ALK is a widely expressed tumour-associated antigen characteristic of ALK	

-positive lymphomas, neuroblastomas and ALK-expressing neoplasias. The peptides of the invention demonstrate cytosolic activity and induce an MHC (major histocompatibility complex) class I restricted cytotoxic CC lymphocyte response against tumor cells expressing the NPM/ALK fusion protein. Hence, the peptides may be utilized during the treatment, via immunotherapy, of ALK-positive lymphomas such as alkomas, neuroblastomas and ALK-expressing neoplasias. In addition, the peptides may be used CC during gene therapy. The current sequence is that of the human p621-629 CC ALK-derived HLA-A\*0201 restricted CTL epitope peptide of the invention XX

Sequence 9 AA;

Query Match 100.0%; Score 42; DB 6; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
1 SLTANKEV 9

RESULT 2  
ABG20950  
ID ABG20950 standard; protein; 675 AA.

AC ABG20950;  
DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #20941.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US008631.

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

PA (HYSE-) HYSEQ INC.

PI Dimaac RT, Liu C, Tang YT;

DR WPI; 2001-633362/73.

DR N-PSDB; AAS85137.

PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensic, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

PS Claim 20; SEQ ID NO 51309; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
XX sequences. (I) is useful as hybridisation probes, polymerase chain  
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
XX and in recombinant production of (II). The polynucleotides are also used  
XX in diagnostics as expressed sequence tags for identifying expressed  
XX genes. (II) is useful in gene therapy techniques to restore normal  
XX activity of (II) or to treat disease states involving (II). (II) is  
XX useful for generating antibodies against it, detecting or quantitating a  
XX polypeptide in tissue, as molecular weight markers and as a food  
XX supplement. (II) and its binding partners are useful in medical imaging  
XX of sites expressing (II). (I) and (II) are useful for treating disorders  
XX involving aberrant protein expression or biological activity. The  
XX polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensic, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pcr\_sequences  
XX

Sequence 675 AA;

Query Match 100.0%; Score 42; DB 4; Length 675;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
616 SLTANKEV 624

RESULT 3  
AAB73858  
ID AAB73858 standard; protein; 680 AA.

AC AAB73858;  
DT 15-MAY-2001 (first entry)

DE Human NPM/ALK fusion protein.

KW Human; anaplastic lymphoma kinase; ALK; nucleophosmin; NPM;  
KW chromosomal translocation; cancer; NPM/ALK; fusion.

OS Homo sapiens.

PN US6174674-B1.

PD 16-JAN-2001.

PF 19-JUN-1998; 98US-00100089.

PR 03-DEC-1993; 93US-00160861.

PR 12-OCT-1995; 95US-00542363.

PA (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.

PI Morris SW, Look AT;

DR WPI; 2001-243208/25.

DR N-PSDB; AAB76868.

PT Detection of chromosomal rearrangement or translocations present in t(2;  
PT 5) for cancer diagnosis comprises using and identifying anaplastic  
PT lymphoma kinase and nucleophosmin genes in nucleic acid hybridization  
PT techniques.

PS Claim 1; Fig 2A; 87pp; English.

XX The present sequence is given in a specification relating to a method for  
XX detecting a chromosomal rearrangement involving a breakpoint in the  
XX anaplastic lymphoma kinase (ALK) or nucleophosmin (NPM) gene. The method  
XX is based on the identification and sequence determination of a novel  
XX gene, ALK, which is fused to the gene encoding NPM in translocations  
XX present in t(2;5) lymphoma cells. The method is useful for detecting a  
XX t(2;5) chromosomal rearrangement, or a chromosomal rearrangement  
XX involving the ALK gene or NPM gene. It is particularly useful in  
XX determining if particular cells or tissues express ALK or NPM/ALK coding  
XX sequences. The method is also useful in diagnostic assays to determine,  
XX for example, if a mammal has cancer or a genetic predisposition to cancer

Sequence 680 AA;

Query Match 100.0%; Score 42; DB 4; Length 680;  
Best Local Similarity 100.0%; Pred. No. 1.6;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLTANKEV 9  
 |||||  
 Db 621 SLTANKEV 629

RESULT 4  
 ABG95067  
 ID ABG95067 standard; protein; 680 AA.  
 XX AC  
 XX ABG95067;  
 XX DT  
 XX 04-DEC-2002 (first entry)  
 XX DE  
 XX Human translocation (2; 5)(p23; q35) protein.  
 XX DE  
 XX Chromosome aberration; oncogenic fusion protein; cancer;  
 XX proliferative disease; cellular protein isoform; heat shock protein 90;  
 XX HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;  
 XX T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;  
 XX acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;  
 XX papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;  
 XX rhabdomyosarcoma; synovial sarcoma; viral infection.  
 XX OS  
 XX Homo sapiens.  
 XX PN  
 XX WO200269900-A2.  
 XX PD  
 XX 12-SEP-2002.  
 XX PF  
 XX 01-MAR-2002; 2002NO-US006518.  
 XX PR  
 XX 01-MAR-2001; 2001US-0272751P.  
 XX PA  
 XX (CONF-) CONFORMA THERAPEUTICS CORP.  
 XX PI  
 XX Fritz LC, Burrows FJ;  
 XX WIPI: 2002-698710/75.  
 XX DR  
 XX N-PSDB; ABS73246.  
 XX PT  
 XX Treating genetically-defined disease associated with chromosomal  
 XX aberrations yielding oncogenic fusion proteins, e.g. cell proliferative  
 XX diseases, involves administering an inhibitor of heat shock protein 90.  
 XX PS  
 XX Disclosure; page 200-202; 389pp; English.  
 XX CC  
 CC The invention describes a method of treating genetically-defined disease  
 CC associated with chromosomal aberrations yielding oncogenic fusion  
 CC proteins (I), treating cancerous cells containing (II) in a heterogeneous  
 CC cell population, treating proliferative diseases associated with mutant  
 CC protein or cellular protein isoforms (II) dependent on heat shock protein  
 CC (HSP)-90, or selectively treating cells expressing (II) involving  
 CC administering HSP90-inhibitor. The method is useful for treating  
 CC genetically-defined disease with chromosomal aberration yielding  
 CC oncogenic fusion protein, treating cancerous cells containing fusion  
 CC protein in heterogeneous cell population, treating proliferative disease  
 CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or  
 CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.  
 CC HSP90), or selectively treating cells expressing mutant protein or cellular  
 CC protein isoform in a patient heterozygous for (II). The method is useful  
 CC for treating a disease e.g. haematopoietic disorder such as T or B cell  
 CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, AML, NHL and CMML,  
 CC or a disease characterised by a solid tumour such as papillary thyroid  
 CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and  
 CC synovial sarcoma. The method is also useful for treating viral  
 CC infections. This represents a protein encoded by the DNA sequence of a  
 CC chromosome aberration  
 XX  
 XX Sequence 680 AA;  
 SQ

Query Match 100.0%; Score 42; DB 5; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 1.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLTANKEV 9  
 |||||  
 Db 621 SLTANKEV 629

RESULT 5  
 ABU04347  
 ID ABU04347 standard; protein; 680 AA.  
 XX AC  
 XX ABU04347;  
 XX DT  
 XX 29-JAN-2003 (first entry)  
 XX DE  
 XX Human expressed protein tag (EPT) #1013.  
 XX DE  
 XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
 XX protease; protease inhibitor; transporter; cytoskeletal protein;  
 XX receptor; transcription factor; cancer; MHC;  
 XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
 XX OS  
 XX Homo sapiens.  
 XX PN  
 XX WO200278524-A2.  
 XX PD  
 XX 10-OCT-2002.  
 XX PF  
 XX 28-MAR-2002; 2002NO-US009671.  
 XX PR  
 XX 28-MAR-2001; 2001US-0279495P.  
 XX PR  
 XX 21-MAY-2001; 2001US-0292544P.  
 XX PR  
 XX 08-AUG-2001; 2001US-0310801P.  
 XX PR  
 XX 01-OCT-2001; 2001US-0326370P.  
 XX PR  
 XX 04-DEC-2001; 2001US-0336780P.  
 XX PR  
 XX 20-FEB-2002; 2002US-0358985P.  
 XX PA  
 XX (ZYCO-) ZYCOs INC.  
 XX PI  
 XX Chiciz RM, Tomlinson AJ, Urban RG;  
 XX WIPI: 2003-040607/03.  
 XX PT  
 XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 XX cytoskeletal proteins, receptors or transcription factors), useful for  
 XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 XX leukaemia.  
 XX PS  
 XX Example 2; SEQ ID NO 1013; 134pp; English.  
 XX CC  
 CC The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide. The purified polypeptide, or the antibody that binds to this  
 CC polypeptide, is useful for treating cancer. The polypeptide is also  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptides and  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
 CC lymphoma or leukaemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: This sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 680 AA;  
 SQ

Sequence 680 AA,

SQ Sequence 680 AA

Sequence 680 AA;

Query Match 100.0%; Score 42; DB 6; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 1.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLTANKEV 9  
 |||||  
 Db 621 SLTANKEV 629

RESULT 8  
 ABG74472  
 ID ABG74472 standard; protein; 680 AA.  
 AC ABG74472;  
 XX  
 DT 11-APR-2003 (first entry)  
 XX  
 DE NPM/ALK fusion construct SEQ ID 4.  
 XX  
 KW ALK; human; nucleophosmin; NPM; anaplastic lymphoma kinase; chromosome 5;  
 KW t(2; 5) lymphoma cell; chromosomal rearrangement; chromosome 2;  
 KW nucleolar phosphoprotein; centromeric; telomeric.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Protein 1..117  
 FT /note="NPM-derived protein fragment"  
 FT Domain 104..115  
 FT /note="potential metal binding domain"  
 FT Protein 118..680  
 FT /note="ALK-derived protein fragment"  
 FT Domain 182..437  
 FT /label= ALK\_catalytic\_domain  
 XX  
 XX US645197-B1.  
 XX  
 PD 17-SEP-2002.  
 XX  
 PF 28-SEP-2000; 2000US-00670827.  
 XX  
 PR 03-DEC-1993; 93US-00160861.  
 PR 12-OCT-1995; 95US-00542363.  
 PR 19-JUN-1998; 98US-00100089.  
 XX  
 FA (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 XX  
 PI Morris SW, Look AT;  
 XX  
 DR WPI: 2003-101739/09.  
 DR N-PSDB; ABQ77107.  
 XX  
 PT Kit for use in method of detecting t(2;5) chromosomal rearrangements or  
 PT rearrangements involving nucleolar phosphoprotein (NPM) or anaplastic  
 PT lymphoma kinase (ALK) comprises probes that hybridizes with NPM and ALK  
 PT genes.  
 XX  
 PS Example 1; Fig 2A; 87pp; English.  
 XX  
 CC This invention describes a novel kit for use in a method of detecting t(2;  
 CC 5) chromosomal rearrangements (CR) or CRs involving nucleolar  
 CC phosphoprotein (NPM) and anaplastic lymphoma kinase (ALK). The method  
 CC comprises labelled probes that hybridise to a sequence of 2442 or 6226  
 CC nucleotides or their complement, and to region of human chromosome 5/2  
 CC having wild-type NPM/ALK genes (the region is centromeric/telomeric to  
 CC nucleotide). This sequence represents a polypeptide sequence described in  
 CC the disclosure of the invention  
 XX  
 SQ Sequence 680 AA;

Query Match 100.0%; Score 42; DB 6; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 1.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLTANKEV 9  
 |||||  
 Db 621 SLTANKEV 629

RESULT 9  
 AAB73857  
 ID AAB73857 standard; protein; 1620 AA.  
 AC AAB73857;  
 XX  
 DT 15-MAY-2001 (first entry)  
 XX  
 DE Human ALK protein.  
 XX  
 KW Human; anaplastic lymphoma kinase; ALK; nucleophosmin; NPM;  
 KW chromosomal translocation; cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6174674-B1.  
 XX  
 PD 16-JAN-2001.  
 XX  
 PF 19-JUN-1998; 98US-00100089.  
 XX  
 PR 03-DEC-1993; 93US-00160861.  
 PR 12-OCT-1995; 95US-00542363.  
 XX  
 PA (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 XX  
 PI Morris SW, Look AT;  
 XX  
 DR WPI: 2001-243208/25.  
 DR N-PSDB; AAF76867.  
 XX  
 PT Detection of chromosomal rearrangement or translocations present in t(2;  
 PT 5) for cancer diagnosis comprises using and identifying anaplastic  
 PT lymphoma kinase and nucleophosmin genes in nucleic acid hybridization  
 PT techniques.  
 XX  
 PS Claim 10; Fig 3B; 87pp; English.  
 XX  
 CC The present sequence is given in a specification relating to a method for  
 CC detecting a chromosomal rearrangement involving a breakpoint in the  
 CC anaplastic lymphoma kinase (ALK) or nucleophosmin (NPM) gene. The method  
 CC is based on the identification and sequence determination of a novel  
 CC gene, ALK, which is fused to the gene encoding NPM in translocations  
 CC present in t(2;5) lymphoma cells. The method is useful for detecting a  
 CC t(2;5) chromosomal rearrangement, or a chromosomal rearrangement  
 CC involving the ALK gene or NPM gene. It is particularly useful in  
 CC determining if particular cells or tissues express ALK or NPM/ALK coding  
 CC sequences. The method is also useful in diagnostic assays to determine,  
 CC for example, if a mammal has cancer or a genetic predisposition to cancer  
 XX  
 SQ Sequence 1620 AA;

Query Match 100.0%; Score 42; DB 4; Length 1620;  
 Best Local Similarity 100.0%; Pred. No. 4.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLTANKEV 9  
 |||||  
 Db 1561 SLTANKEV 1569

RESULT 10  
 AAO18498  
 ID AAO18498 standard; protein; 1620 AA.  
 AC AAO18498;  
 XX

DT 11-OCT-2002 (first entry)  
 DE Human insulin receptor signaling modifier SEQ ID NO: 18.  
 XX Human, insulin receptor signaling; insulin receptor signaling modifier;  
 KM ISM; diabetes; metabolic syndrome; antidiabetic.  
 XX Homo sapiens.  
 OS  
 XX WO200255664-A2.  
 PN  
 XX  
 PD 18-JUL-2002.  
 XX  
 PF 11-JAN-2002; 2002WO-US001048.  
 XX  
 PR 12-JAN-2001; 2001US-0261226P.  
 PR 12-JAN-2001; 2001US-0261303P.  
 PR 12-JAN-2001; 2001US-0261304P.  
 PR 12-JAN-2001; 2001US-0261335P.  
 PR 12-JAN-2001; 2001US-0261336P.  
 PR 12-JAN-2001; 2001US-0261361P.  
 PR 12-JAN-2001; 2001US-0261456P.  
 PR 12-JAN-2001; 2001US-0261457P.  
 PR 12-JAN-2001; 2001US-0261458P.  
 PR 12-JAN-2001; 2001US-0261459P.  
 PR 12-JAN-2001; 2001US-0261461P.  
 PR 12-JAN-2001; 2001US-0261518P.  
 PR 12-JAN-2001; 2001US-0261531P.  
 PR 12-JAN-2001; 2001US-0261532P.  
 PR 12-JAN-2001; 2001US-0261589P.  
 PR 12-JAN-2001; 2001US-0261590P.  
 PR 12-JAN-2001; 2001US-0261694P.  
 PR 12-JAN-2001; 2001US-0261695P.  
 PR 12-JAN-2001; 2001US-0261697P.  
 XX  
 XX (EXEL-) EXELIXIS INC.  
 XX  
 XX Seidel-Dugan C, Ferguson KC, Kidd T;  
 PI WPI; 2002-599664/64.  
 XX  
 DR N-PDB; AAL48617.  
 XX  
 PT Identifying an insulin receptor signaling modulator, useful as drug  
 PT targets for treating diabetes or metabolic disorders, comprises  
 PT contacting an assay system comprising insulin receptor signaling  
 PT modifiers with a test agent.  
 XX  
 PS Disclosure; Page 59-66; 232pp; English.  
 XX  
 CC The present invention relates to a method of identifying a candidate  
 CC insulin receptor (INR) signaling modulating agent, involving contacting  
 CC an assay system comprising an insulin receptor signaling modifier (ISM)  
 CC polypeptide or nucleic acid with a test agent, and detecting a test agent  
 CC -biased activity of the assay system. The method is useful for  
 CC identifying candidate INR signaling modulating agents. ISM genes may be  
 CC used as drug targets for treatment of disorders related to INR signaling  
 CC such as diabetes or metabolic syndrome. ISM nucleic acids and  
 CC polypeptides are useful for identifying and testing agents that modulate  
 CC INR function and for other applications related to the involvement of ISM  
 CC in INR signaling, and for identifying subjects having a predisposition to  
 CC such diseases associated with INR signaling. The present sequence is an  
 CC ISM protein described in the exemplification of the invention  
 XX  
 SQ Sequence 1620 AA;  
 XX  
 QY Query Match 100.0%; Score 42; DB 5; Length 1620;  
 DB Best Local Similarity 100.0%; Pred. No. 4.e';  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0,  
 1 STRANKEV 9  
 |||||  
 1561 SLTANKEV 1569

```

RESULT 11
ID ABG74471
XX ABG74471 standard; protein; 1620 AA.
AC ABG74471;
XX
DT 11-APR-2003 (first entry)
XX
DE Human ALK protein SEQ ID 2.
XX
KW ALK; human; nucleophosmin; NPM; anaplastic lymphoma kinase; chromosome 5;
XX t(2; 5) lymphoma cell; chromosomal rearrangement; chromosome 2;
XX nucleolar phosphoprotein; centromeric; telomeric.
XX
OS Homo sapiens.
XX
FH Location/Qualifiers
FT 1..26
FT /label=signal_peptide
FT 27..1620
FT /label=mature_ALK
FT 1031..1058
FT /label=transmembrane_domain
FT 1058..1059
FT /note="NPM-ALK fusion junction"
FT 1123..1377
FT /label=tyrosine_kinase_catalytic_domain
XX
XX US6451997-B1.
XX
PD 17-SEP-2002.
XX
XX 28-SEP-2000; 2000US-00670827.
XX
XX 03-DEC-1993; 93US-00160861.
XX PR 12-OCT-1995; 95US-00542853.
XX PR 19-JUN-1998; 98US-00100089.
XX
XX (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
XX Morris SW, Look AT;
XX
XX WPI: 2003-101739/09.
XX DR N-PSDB; ABG77106.
XX
XX Kit for use in method of detecting t(2;5) chromosomal rearrangements or
XX rearrangements involving nucleolar phosphoprotein (NPM) or anaplastic
XX lymphoma kinase (ALK) comprises probes that hybridizes with NPM and ALK
XX genes.
XX
XX Example 1; Fig 3B; 87bp; English.
XX
XX This invention describes a novel kit for use in a method of detecting t(2
XX 5) chromosomal rearrangements (CR) or CRs involving nucleolar
XX phosphoprotein (NPM) and anaplastic lymphoma kinase (ALK). The method
XX comprises labeled probes that hybridize to a sequence of 2442 or 6226
XX nucleotides or their complement, and to region of human chromosome 5/2
XX having wild-type NPM/ALK genes (the region is centromeric/telomeric to
XX nucleotide). This sequence represents a polypeptide sequence described in
XX the disclosure of the invention
XX
XX Sequence 1620 AA;
XX
XX Query Match 100.0%; Score 42; DB 6; Length 1620;
XX Best Local Similarity 100.0%; Pred. NO. 4.6;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 STJANKKEY 9
XX |||||||
DB 1561 STJANKKEY 1569

```

RESULT 12  
 AAR70013 standard; protein; 660 AA.  
 XX  
 AC AAR70013;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 25-SEP-1995 (first entry)  
 XX  
 DE tick carboxypeptidase.  
 XX  
 KM tick; antigen; carboxypeptidase; vaccine.  
 XX  
 OS Boophilus microplus.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..29  
 FT /label= signal  
 FT 30..53  
 FT /label= BM91 peptide  
 FT 185..201  
 FT /label= peptide AAT91251  
 FT 202..209  
 FT /label= AAT91141  
 FT 225..253  
 FT /label= T9118  
 FT 262..269  
 FT /label= Bm91 peptide  
 FT 442..452  
 FT /label= T9129(a)  
 FT 456..481  
 FT /label= T9129(b)  
 FT /label= T9109  
 FT 559..606  
 FT /label= T9109  
 FT 639..655  
 FT /label= C-terminal transmembrane domain  
 FT /note= "putative"  
 XX  
 PN WO9504827-A1.  
 XX  
 PD 16-FEB-1995.  
 XX  
 PF 10-AUG-1994; 94WO-AU000463.  
 XX  
 PR 10-AUG-1993; 93AU-00000458.  
 XX  
 PA (BIOT-) BIOTECH AUSTRALIA PTY LTD.  
 PA (CSIR) COMMONWEALTH SCI & IND RES ORG.  
 XX  
 PI Cobon GS, Willadeen P, Kemp DH, Tellam RL;  
 XX  
 XX WPI, 1995-090905/12.  
 DR N-PSDB; AAO82948.  
 XX  
 PT New antigenic tick carboxypeptidase and corresp. DNA - are used in  
 PT vaccines for producing antibodies against ticks, insects and nematodes.  
 XX  
 PS Disclosure; Fig 6; 138pp; English.  
 XX  
 CC Clone A5 was prep'd. from adult tick cDNA library. Clone 4U1 was prep'd.  
 CC from the larval stage of B. microplus (Calliope strain). AAO82948 is a  
 CC hybrid of sequences from clone 4U1 (nt 1-966 & 1747-2047) and A5 (nts 967  
 CC -1746). The translation of the tick carboxypeptidase cDNA sequences is  
 CC shown in AAR70013. All the native tick carboxypeptidase sequences listed  
 CC in Table 11 (see AAR70014-R70023) are found in the translation (see FT).  
 CC The predicted A5 sequence agrees with the peptide sequence for all  
 CC peptides with 2 exceptions. These differences are Asp for Glu14 in  
 CC peptide T9126, and Asn for Asp12 in peptide T9118. Tick carboxypeptidase  
 CC has a moi. wt. 75172. In addition to the features in FT, the  
 CC carboxypeptidase A5 sequence has 8 potential N-linked glycosylation sites  
 CC and a potential glycosylphosphatidyl inositol anchor sequence similar to  
 CC that found in Bm86. It has significant homology with zinc dependent  
 CC dipeptidyl carboxypeptidases from mammals. (Updated on 25-MAR-2003 to

CC correct PN field.)  
 XX  
 XX Sequence 660 AA;  
 SQ  
 Query Match 78.6%; Score 33; DB 2; Length 660;  
 Beef Local Similarity 87.5%; Pred. No. 1.8e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 LTRANKREV 9  
 |||||  
 Db 173 LTRANKREV 180

RESULT 13  
 AAY05901  
 ID AAY05901 standard; protein; 510 AA.  
 XX  
 AC AAY05901;  
 XX  
 DT 02-AUG-1999 (first entry)  
 XX  
 DE Tobacco fatty acid hydroxylase CYP94A6.  
 XX  
 KM Fatty acid omega-hydroxylase; cytochrome P450; transgenic plant; lipid;  
 KM hydroxylation; epoxidation; oilseed; vegetable oil; crop protection;  
 KM omega-hydroxy acid; CYP94A6; tobacco.  
 XX  
 OS Nicotiana tabacum.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 317..328  
 FT /note= "haem-binding domain, corresponds to signature  
 FT motif claimed in Claim 2"  
 XX  
 PN WO9318224-A1.  
 XX  
 PD 15-APR-1999.  
 XX  
 PF 06-OCT-1998; 98WO-IB001716.  
 XX  
 PR 06-OCT-1997; 97US-0060960P.  
 XX  
 PA (CNRS) CENT NAT RECH SCI.  
 XX  
 PI Tillet N, Pinot F, Benveniste I, Le Bouquin R, Helyig C, Batard Y;  
 PI Cabellac-Huarrado F, Werck-Reichhart D, Salaun J, Durst F;  
 XX  
 DR WPI, 1999-264030/22.  
 DR N-PSDB; AAX58405.  
 XX  
 PT Nucleic acid encoding plant fatty acid hydroxylases.  
 XX  
 PS Example 7; Fig 26A-B; 157pp; English.  
 XX  
 CC This sequence represents a cytochrome P450 protein, CYP94A6, of tobacco  
 CC that is expected to have fatty acid hydroxylase activity since it  
 CC displays the characteristic signature motif for this class of enzymes.  
 CC The invention provides isolated nucleic acids (see AAX58400-06) encoding  
 CC plant fatty acid hydroxylases (see AAY05886-902). Also claimed are host  
 CC cells, transgenic plants and compositions consisting of the plant fatty  
 CC acid hydroxylase, a process for isolating additional fatty acid  
 CC composition in a plant by expressing the plant fatty acid hydroxylase in  
 CC a transgenic plant, and hydroxylating or epoxidating a fatty acid  
 CC substrate in the plant. Manipulating the hydroxylated fatty acid content  
 CC of plants will modify resistance to drought and attack by insects and  
 CC other pests. The transgenic plants may also be used as sources of  
 CC hydroxylated and epoxidized fatty acids useful in the manufacture of e.g.  
 CC lubricants, anti-slip agents, plasticizers, coating agents, detergents  
 CC and surfactants  
 XX  
 SQ Sequence 510 AA;

Query Match 76.2%; Score 32; DB 2; Length 510;  
 Best Local Similarity 87.5%; Pred. No. 2.2e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LTANKEV 9  
 DB 494 LTANKEV 501

RESULT 14  
 ID AB047568  
 XX AB047568 standard; protein; 541 AA.

AC AB047568;  
 XX

DT 05-FEB-2002 (first entry)

DE Listeria monocytogenes protein #272.

KM Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;  
 KM vitamin B12; bacterial infection; disease.

OS Listeria monocytogenes.

PN WO200177335-A2.

PD 18-OCT-2001.

PF 11-APR-2001; 2001WO-FR001118.

PR 11-APR-2000; 2000FR-00004629.

XX (INSP) INST PASTEUR.

XX Buehrleier C, Frangeul L, Couve E, Rusnick C, Feihl H, Dehoux P,  
 PI Dusunget O, Cretouan F, Nedjati H, Glaser P, Cossart P,  
 PI Damiens J, Goebel W, Xieft J, Kuhn M, Ng E, Vazquez-Boland JA,  
 PI Dominguez-Bernal G, Garrido-Garcia P, Tixeront-Martinez A, Amend A,  
 PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L,  
 PI Perez-Diaz J, Baguerio F, Garcia Del Portillo F, Gomez-Lopez N,  
 PI Madueno E, De Pablos B, Wehlend J, Kaerst U, Entian K, Hauf J,  
 PI Rose M, Voss H;  
 XX WPI: 2002-010914/01.

PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment  
 PT and prevention of Listeria and related bacterial infections, and related  
 PT polypeptides.

PS Claim 6; SEQ ID NO 273; 192bp; French.

CC The present invention relates to the genome sequence of Listeria  
 CC monocytogenes EGD-e (see AB03041). The genome sequence and fragments of  
 CC it are useful for selecting probes and primers for detecting genes in L.  
 CC monocytogenes and related organisms, and for studying genetic  
 CC polymorphisms and other genomes. The present sequence is a protein  
 CC encoded by the genome sequence of the present invention. Proteins  
 CC expressed from the genome sequence are useful for raising specific  
 CC antibodies. Identification of L. monocytogenes and related organisms, and  
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin  
 CC B12. The genome sequence and proteins encoded by it are also useful for  
 CC selecting compounds that regulate gene expression and cell replication  
 CC and modulate L. monocytogenes-related diseases. In addition, the genome  
 CC sequence and proteins encoded by it are useful in pharmaceutical and  
 CC vaccine compositions for the treatment or prevention of infections by L.  
 CC monocytogenes and related organisms. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp://ipo.int/pub/published\_pcr\_sequences

XX Sequence 541 AA;

Query Match 76.2%; Score 32; DB 5; Length 541;

Best Local Similarity 66.7%; Pred. No. 2.3e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SFTANKEV 9  
 DB 154 SFTANKEV 162

RESULT 15  
 ID AB032622  
 XX AB032622 standard; protein; 541 AA.

AC AB032622;  
 XX

DT 19-JUN-2003 (first entry)

DE Protein encoded by prokaryotic essential gene #8149.

KM Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Listeria monocytogenes.

PN WO200277183-A2.

PD 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362659P.

XX (ELIT) ELITRA PHARM INC.  
 XX Wang L, Zamudio C, Malone C, Hasebeck R, Ohlsen XL, Zyskind JW,  
 PI Wall D, Traxler JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 DR WPI: 2003-023926/02.  
 XX N-PSDB; ACR36492.

PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.

PS Claim 25; SEQ ID NO 60546; 176bp; English.

CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or its gene product lies  
 CC pathway in which a proliferation-regulated gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids

CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX

SO Sequence 541 AA:

Query Match Best Local Similarity 76.2%; Score 32; DB 6; Length 541;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
 Db 154 SFVANMKEI 162

Search completed: November 30, 2004, 07:39:59  
 Job time : 71.715 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: November 30, 2004, 07:19:22 ; Search time 339.231 Seconds  
(without alignments)  
9.410 Million cell updates/sec

Title: US-10-008-377A-7

Perfect score: 42

Sequence: 1 SLTANKREV 9

## Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1575965 seqs, 354694765 residues

Total number of hits satisfying chosen parameters: 1575965

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpa/US07\_PUBCOMB.pep:\*

2: /cgn2\_6/ptodata/1/pubpa/PTC\_NEW\_PUB.pep:\*

3: /cgn2\_6/ptodata/1/pubpa/US06\_NEW\_PUB.pep:\*

4: /cgn2\_6/ptodata/1/pubpa/US06\_PUBCOMB.pep:\*

5: /cgn2\_6/ptodata/1/pubpa/US07\_NEW\_PUB.pep:\*

6: /cgn2\_6/ptodata/1/pubpa/PTCUS\_PUBCOMB.pep:\*

7: /cgn2\_6/ptodata/1/pubpa/US08\_NEW\_PUB.pep:\*

8: /cgn2\_6/ptodata/1/pubpa/US08\_PUBCOMB.pep:\*

9: /cgn2\_6/ptodata/1/pubpa/US09\_PUBCOMB.pep:\*

10: /cgn2\_6/ptodata/1/pubpa/US09\_PUBCOMB.pep:\*

11: /cgn2\_6/ptodata/1/pubpa/US09C\_PUBCOMB.pep:\*

12: /cgn2\_6/ptodata/1/pubpa/US09\_NEW\_PUB.pep:\*

13: /cgn2\_6/ptodata/1/pubpa/US10\_PUBCOMB.pep:\*

14: /cgn2\_6/ptodata/1/pubpa/US10\_PUBCOMB.pep:\*

15: /cgn2\_6/ptodata/1/pubpa/US10C\_PUBCOMB.pep:\*

16: /cgn2\_6/ptodata/1/pubpa/US10C\_PUBCOMB.pep:\*

17: /cgn2\_6/ptodata/1/pubpa/US10C\_PUBCOMB.pep:\*

18: /cgn2\_6/ptodata/1/pubpa/US10C\_PUBCOMB.pep:\*

19: /cgn2\_6/ptodata/1/pubpa/US60\_NEW\_PUB.pep:\*

20: /cgn2\_6/ptodata/1/pubpa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	42	100.0	9	US-10-008-377-7
2	42	100.0	680	US-09-827-949-4
3	42	100.0	1620	US-09-827-949-2
4	35	83.3	268	US-10-424-599-202343
5	33	76.6	96	US-10-425-115-343624
6	33	76.6	1239	US-10-437-963-149396
7	32	76.2	210	US-10-437-963-195871
8	32	76.2	541	US-10-282-122A-60546
9	31	73.8	41	US-10-282-115-873
10	31	73.8	116	US-10-425-114-43189
11	31	73.8	162	US-10-425-115-242109
12	31	73.8	202	US-10-424-599-211762
13	31	73.8	339	US-10-282-122A-53169

14	31	73.8	367	US-10-424-599-178778	Sequence 178778, A
15	31	73.8	510	US-10-282-122A-60559	Sequence 60559, A
16	31	73.8	1594	US-10-425-115-291526	Sequence 291526, A
17	31	73.8	2545	US-10-042-865-12	Sequence 12, Appl
18	30	71.4	179	US-10-283-122A-71093	Sequence 71093, A
19	30	71.4	285	US-10-365-493-16391	Sequence 16391, A
20	30	71.4	297	US-10-425-115-425348	Sequence 225348, A
21	30	71.4	503	US-09-871-212-7	Sequence 7, Appl
22	30	71.4	611	US-10-282-122A-71118	Sequence 71118, A
23	30	71.4	608	US-10-381-779-30	Sequence 30, Appl
24	30	71.4	608	US-10-369-493-1058	Sequence 3058, Ap
25	30	71.4	868	US-10-437-963-123248	Sequence 123248, A
26	30	71.4	1082	US-10-437-963-167162	Sequence 167162, A
27	30	71.4	1734	US-09-852-027-81	Sequence 81, Appl
28	30	71.4	1734	US-10-377-035-14	Sequence 14, Appl
29	30	71.4	1734	US-10-042-865-52	Sequence 52, Appl
30	30	71.4	1798	US-10-311-034-23	Sequence 23, Appl
31	30	71.4	1798	US-10-618-941-70	Sequence 70, Appl
32	30	71.4	1805	US-10-425-114-54369	Sequence 54369, A
33	29	69.0	68	US-10-425-115-232919	Sequence 232919, A
34	29	69.0	147	US-10-767-701-58957	Sequence 58957, A
35	29	69.0	154	US-10-425-115-293462	Sequence 293462, A
36	29	69.0	159	US-10-767-701-62385	Sequence 62385, A
37	29	69.0	178	US-10-424-599-167198	Sequence 167198, A
38	29	69.0	202	US-10-767-701-42169	Sequence 42169, A
39	29	69.0	222	US-10-312-273-339	Sequence 339, App
40	29	69.0	222	US-10-289-762-783	Sequence 783, App
41	29	69.0	283	US-10-424-599-776594	Sequence 276904, A
42	29	69.0	334	US-10-176-306-38	Sequence 38, Appl
43	29	69.0	342	US-10-767-701-44976	Sequence 44976, A
44	29	69.0	392	US-10-282-122A-74587	Sequence 74587, A
45	29	69.0	482	US-10-365-493-1819	Sequence 1819, Ap

## ALIGNMENTS

RESULT 1

US-10-008-377-7

; Sequence 7, Application US/10008377

; Publication No. US20030157101A1

; GENERAL INFORMATION:

; APPLICANT: Gambacorti-Passerini, Carlo

; APPLICANT: Passoni, Lorena

; TITLE OF INVENTION: Immunogenic Alk Peptides

; FILE REFERENCE: 045922/241203

; CURRENT APPLICATION NUMBER: US/10/008,377

; CURRENT FILING DATE: 2001-11-15

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: FASTSEQ for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-008-377-7

Query Match 100.0%; Score 42; DB 14; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.4e+06;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKREV 9

DB 1 SLTANKREV 9

RESULT 2

US-09-827-949-4

; Sequence 4, Application US/09827949

; Patent No. US20010021505A1

; GENERAL INFORMATION:

; APPLICANT: Morris, Stephan W.

; APPLICANT: Look, A. Thomas

; TITLE OF INVENTION: Alk Protein Tyrosine Kinase/Receptor and Ligands Thereof

FILE REFERENCE: 0656.0400004  
CURRENT APPLICATION NUMBER: US/09/827,949  
CURRENT FILING DATE: 2001-04-09  
PRIOR APPLICATION NUMBER: US 09/570,827  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 09/100,089  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: US 08/542,363  
PRIOR FILING DATE: 1995-10-12  
PRIOR APPLICATION NUMBER: US 08/160,861  
PRIOR FILING DATE: 1993-12-03  
NUMBER OF SEQ ID NOS: 43  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 4  
LENGTH: 680  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-827-949-4

Query Match 100.0%; Score 42; DB 9; Length 680;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 SLTANKEV 9  
DB 621 SLTANKEV 629

RESULT 3  
US-09-827-949-2  
Sequence 2, Application US/09827949  
Patent No. US20010021505A1  
GENERAL INFORMATION:  
APPLICANT: Morrie, Stephan W.  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof  
FILE REFERENCE: 0656.0400004  
CURRENT APPLICATION NUMBER: US/09/827,949  
CURRENT FILING DATE: 2001-04-09  
PRIOR APPLICATION NUMBER: US 09/570,827  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 09/100,089  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: US 08/542,363  
PRIOR FILING DATE: 1995-10-12  
PRIOR APPLICATION NUMBER: US 08/160,861  
PRIOR FILING DATE: 1993-12-03  
NUMBER OF SEQ ID NOS: 43  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 2  
LENGTH: 1620  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-827-949-2

Query Match 100.0%; Score 42; DB 9; Length 1620;  
Best Local Similarity 100.0%; Pred. No. 9.7;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 SLTANKEV 9  
DB 1561 SLTANKEV 1569

RESULT 4  
US-10-424-599-202343  
Sequence 202343, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 202343  
LENGTH: 268  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_24740C.1.pap  
US-10-424-599-202343

Query Match 83.3%; Score 35; DB 15; Length 268;  
Best Local Similarity 77.8%; Pred. No. 33;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

CY 1 SLTANKEV 9  
DB 175 SLTANKEV 183

RESULT 5  
US-10-425-115-343624  
Sequence 343624, Application US/10425115  
Publication No. US20040214272A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants  
FILE REFERENCE: 38-21(53222)B  
CURRENT APPLICATION NUMBER: US/10/425,115  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 343624  
LENGTH: 96  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: MRT4577\_76547C.1.pap  
US-10-425-115-343624

Query Match 78.6%; Score 33; DB 17; Length 96;  
Best Local Similarity 66.7%; Pred. No. 33;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

CY 1 SLTANKEV 9  
DB 13 SLTANKEV 21

RESULT 6  
US-10-437-963-149396  
Sequence 149396, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barzduk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14

```

; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 149396
; LENGTH: 1239
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) - (1239)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_49733C.1.pep
US-10-437-963-149396

Query Match
Best Local Similarity 78.6%; Score 33; DB 16; Length 1239;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKE 8
DB 519 SLTANKE 526

RESULT 7
US-10-437-963-195871
; Sequence 195871, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Bouharov, Andrey A.
; APPLICANT: Barzak, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21 (53221) B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 195871
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91778C.1.pep
US-10-437-963-195871

Query Match
Best Local Similarity 76.2%; Score 32; DB 16; Length 210;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9
DB 7 SLTANKEV 15

RESULT 8
US-10-282-122A-60546
; Sequence 60546, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malore, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Olsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
```

```

; APPLICANT: Foreyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent version 3.1
; SEQ ID NO 60546
; LENGTH: 541
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-10-282-122A-60546

Query Match
Best Local Similarity 76.2%; Score 32; DB 15; Length 541;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTANKEV 9
DB 154 SLTANKEV 162

RESULT 9
US-10-296-115-873
; Sequence 873, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseg Inc
; TITLE OF INVENTION: No. US20040053248A1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 873
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-296-115-873

Query Match
Best Local Similarity 73.8%; Score 31; DB 15; Length 41;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9
DB 12 SLTANKEV 20
```

RESULT 10

US-10-425-114-43189  
; Sequence 43189, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 43189  
; LENGTH: 116  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700748151\_FLI.pep  
US-10-425-114-43189

Query Match  
Best Local Similarity 73.8%; Score 31; DB 15; Length 116;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLTANMKE 8  
|||  
DB 77 SLTANMKE 84

RESULT 11

US-10-425-115-242109  
; Sequence 242109, Application US/10425115  
; Publication No. US2004021472A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(5322)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 242109  
; LENGTH: 162  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(162)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MWT4577\_152386C.1.pep  
US-10-425-115-242109

Query Match  
Best Local Similarity 73.8%; Score 31; DB 17; Length 162;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLTANMKEV 9  
|||  
DB 116 SLTANMKEV 124

RESULT 12  
US-10-424-599-211762

RESULT 13

US-10-282-122A-53169  
; Sequence 53169, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trewick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
US-10-282-122A-53169

Query Match  
Best Local Similarity 73.8%; Score 31; DB 15; Length 202;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLTANMKE 8  
|||  
DB 163 SLTANMKE 170

PRIOR APPLICATION NUMBER: 60/269,308  
 PRIOR FILING DATE: 2001-02-16  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 78614  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 53169  
 LENGTH: 339  
 TYPE: PRT  
 ORGANISM: Clostridium difficile  
 US-10-282-122A-53169

Query Match 73.8%; Score 31; DB 15; Length 339;  
 Best Local Similarity 66.7%; Pred. No. 3.6e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLTANKKEV 9  
 ||:|||||  
 Db 98 SLTANKKEV 106

RESULT 14  
 US-10-424-599-178778  
 Sequence 178778, Application US/10424599  
 Publication No. US20040031072A1  
 GENERAL INFORMATION:  
 APPLICANT: La Rosa, Thomas J  
 APPLICANT: Kovacic David K  
 APPLICANT: Zhou Yihua  
 APPLICANT: Cao Yongwei  
 TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(53223)B  
 CURRENT APPLICATION NUMBER: US/10/424,599  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 285664  
 SEQ ID NO 178778  
 LENGTH: 367  
 TYPE: PRT  
 ORGANISM: Glycine max  
 FEATURE:  
 NAME/KEY: unsure  
 LOCATION: (1)-(367)  
 OTHER INFORMATION: unsure at all Xaa locations  
 FEATURE:  
 OTHER INFORMATION: Clone ID: PAT\_MRT3847\_132453C.1.Pep  
 US-10-424-599-178778

Query Match 73.8%; Score 31; DB 15; Length 367;  
 Best Local Similarity 75.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKKE 8  
 ||:|||||  
 Db 305 SLTANKKE 312

RESULT 15  
 US-10-282-122A-60559  
 Sequence 60559, Application US/10282122A  
 Publication No. US20040029129A1  
 GENERAL INFORMATION:  
 APPLICANT: Wang, Liangsu  
 APPLICANT: Zamudio, Carlos  
 APPLICANT: Malone, Cheryl  
 APPLICANT: Haseibeck, Robert  
 APPLICANT: Ohlsen, Kari  
 APPLICANT: Zyskind, Judith  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John  
 APPLICANT: Carr, Grant  
 APPLICANT: Yamamoto, Robert  
 APPLICANT: Forsyth, R.  
 APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 FILE REFERENCE: EPIGRA.034A  
 CURRENT APPLICATION NUMBER: US/10/282,122A  
 CURRENT FILING DATE: 2003-02-20  
 PRIOR APPLICATION NUMBER: 60/191,078  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/206,848  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: 60/230,335  
 PRIOR FILING DATE: 2000-09-06  
 PRIOR APPLICATION NUMBER: 60/230,347  
 PRIOR FILING DATE: 2000-09-09  
 PRIOR APPLICATION NUMBER: 60/242,578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/267,636  
 PRIOR FILING DATE: 2001-02-09  
 PRIOR APPLICATION NUMBER: 60/269,308  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 78614  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 60559  
 LENGTH: 510  
 TYPE: PRT  
 ORGANISM: Listeria monocytogenes  
 US-10-282-122A-60559

Query Match 73.8%; Score 31; DB 15; Length 510;  
 Best Local Similarity 55.6%; Pred. No. 5.7e+02;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKKEV 9  
 ||:|||||  
 Db 224 TITANKKDL 232

Search completed: November 30, 2004, 08:06:37  
 Job time : 340.231 secs

**This Page Blank (uspio)**

Tue Nov 30 08:49:32 2004

us-10-008-377a-7.open.ra1

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:12:12 (Search time 22.6154 Seconds  
(without alignments)  
26.392 Million cell updates/sec

Title: US-10-008-377A-7  
Perfect score: 42  
Sequence: 1 SLTANKKEV 9

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues  
Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Issued Patents RA: \*  
1: /cgn2\_6/prodata/1/1aa/5A.CONB.pep: \*  
2: /cgn2\_6/prodata/1/1aa/5B.CONB.pep: \*  
3: /cgn2\_6/prodata/1/1aa/6A.CONB.pep: \*  
4: /cgn2\_6/prodata/1/1aa/6B.CONB.pep: \*  
5: /cgn2\_6/prodata/1/1aa/PCITUS.CONB.pep: \*  
6: /cgn2\_6/prodata/1/1aa/backfiles.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	42	100.0	US-08-543-363-4	Sequence 4, App1
2	42	100.0	US-09-100-089-4	Sequence 4, App1
3	42	100.0	US-09-670-827-4	Sequence 4, App1
4	42	100.0	US-09-827-949-4	Sequence 4, App1
5	42	100.0	US-08-542-363-2	Sequence 2, App1
6	42	100.0	US-09-100-089-2	Sequence 2, App1
7	42	100.0	US-09-670-827-2	Sequence 2, App1
8	42	100.0	US-09-827-949-2	Sequence 2, App1
9	31	73.8	US-09-134-000C-6610	Sequence 6610, Ap
10	30	71.4	US-09-710-279-1808	Sequence 1808, Ap
11	30	71.4	US-09-134-001C-4215	Sequence 4215, Ap
12	30	71.4	US-09-583-110-3178	Sequence 3178, Ap
13	30	71.4	US-09-583-110-4472	Sequence 4472, Ap
14	30	71.4	US-09-252-991A-16966	Sequence 16966, A
15	30	71.4	US-09-710-279-320	Sequence 320, App
16	30	71.4	US-09-328-352-5957	Sequence 5957, Ap
17	30	71.4	US-09-134-000C-4057	Sequence 4057, Ap
18	30	71.4	US-09-710-279-232	Sequence 232, App
19	30	71.4	US-09-248-796A-14512	Sequence 14512, A
20	30	71.4	US-09-134-001C-4248	Sequence 4248, Ap
21	29	69.0	US-09-198-452A-783	Sequence 783, App
22	29	69.0	US-09-328-352-4901	Sequence 4901, Ap
23	29	69.0	US-09-543-681A-7802	Sequence 7802, Ap
24	29	69.0	US-09-071-035-364	Sequence 364, App
25	29	69.0	US-09-370-838-67	Sequence 67, App1
26	29	69.0	US-09-071-035-362	Sequence 362, App
27	29	69.0	US-09-538-092-944	Sequence 944, App

28	29	69.0	764	4	US-09-854-133-67	Sequence 67, App1
29	29	69.0	776	4	US-09-134-000C-5717	Sequence 5717, Ap
30	28	66.7	100	4	US-09-248-796A-26482	Sequence 26482, A
31	28	66.7	104	3	US-08-946-329A-97	Sequence 97, App1
32	28	66.7	104	4	US-09-248-796A-25322	Sequence 25322, A
33	28	66.7	156	4	US-09-540-236-2381	Sequence 2381, Ap
34	28	66.7	189	4	US-09-270-767-35780	Sequence 35780, A
35	28	66.7	189	4	US-09-270-767-50997	Sequence 50997, A
36	28	66.7	251	4	US-09-328-352-4867	Sequence 4867, Ap
37	28	66.7	393	4	US-09-248-796A-18084	Sequence 18084, A
38	28	66.7	510	4	US-09-134-000C-5774	Sequence 5774, Ap
39	28	66.7	521	4	US-09-107-532A-6431	Sequence 6431, Ap
40	28	66.7	709	1	US-07-814-964-7	Sequence 7, App1
41	28	66.7	709	1	US-08-258-442-7	Sequence 7, App1
42	28	66.7	709	1	US-08-328-809-2	Sequence 2, App1
43	28	66.7	709	3	US-09-015-003-2	Sequence 2, App1
44	28	66.7	709	4	US-08-866-840-2	Sequence 2, App1
45	28	66.7	709	4	US-09-538-092-1287	Sequence 1287, Ap

#### ALIGNMENTS

RESULT 1  
US-08-542-363-4  
Sequence 4, Application US/08542363  
Patent No. 5770421  
GENERAL INFORMATION:  
APPLICANT: MORRIS, Stephan W.  
APPLICANT: LOOK, A. Thomas  
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and  
TITLE OF INVENTION: Ligands thereof  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESS: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/542,363  
FILING DATE: 12-Oct-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.040001/SLF/GKT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2660  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ. ID NO. 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 680 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-542-363-4

Query Match 100.0% Score 42; DB 1; Length 680;  
Best Local Similarity 100.0% Pred. No. 0.44; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKKEV 9  
DB 621 SLTANKKEV 629

## RESULT 2

US-09-100-089-4

Sequence 4, Application US/09100089

Patent No. 6174674

GENERAL INFORMATION:

APPLICANT: Morris, Stephan W.

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and

TITLE OF INVENTION: Ligands Thereof

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESS: STERN, KESSLER, GOLDSTEIN &amp; FOX P.L.L.C.

STREET: 1100 New York Avenue, N.W., Suite 600

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/100,089

FILING DATE: Herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/160,861

FILING DATE: 03-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/542,363

FILING DATE: 12-OCT-1995

ATTORNEY/AGENT INFORMATION:

NAME: Fox, Samuel L.

REGISTRATION NUMBER: 30,353

REFERENCE/DOCKET NUMBER: 0656.0400002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 680 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-100-089-4

Query Match

Best Local Similarity 100.0%; Score 42; DB 3; Length 680;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9

Db 621 SLTANKEV 629

## RESULT 3

US-09-670-827-4

Sequence 4, Application US/09670827

Patent No. 6451997

GENERAL INFORMATION:

APPLICANT: Morris, Stephan W.

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and

TITLE OF INVENTION: Ligands Thereof

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESS: STERN, KESSLER, GOLDSTEIN &amp; FOX P.L.L.C.

STREET: 1100 New York Avenue, N.W., Suite 600

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/670,827

FILING DATE: 28-SEP-2000

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/160,861

FILING DATE: 03-DEC-1993

APPLICATION NUMBER: US 08/542,363

FILING DATE: 12-OCT-1995

APPLICATION NUMBER: US 09/100,089

FILING DATE: 19-JUN-1998

ATTORNEY/AGENT INFORMATION:

NAME: Fox, Samuel L.

REGISTRATION NUMBER: 30,353

REFERENCE/DOCKET NUMBER: 0656.0400003

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 680 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-670-827-4

Query Match 100.0%; Score 42; DB 4; Length 680;

Best Local Similarity 100.0%; Pred. No. 0.44;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9

Db 621 SLTANKEV 629

## RESULT 4

US-09-827-949-4

Sequence 4, Application US/09827949

Patent No. 6685548

GENERAL INFORMATION:

APPLICANT: Morris, Stephan W.

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof

FILE REFERENCE: 0656.0400004

CURRENT APPLICATION NUMBER: US/09/827,949

CURRENT FILING DATE: 2001-04-09

PRIOR APPLICATION NUMBER: US 09/670,827

PRIOR FILING DATE: 2000-09-28

PRIOR APPLICATION NUMBER: US 09/100,089

PRIOR FILING DATE: 1998-06-19

PRIOR APPLICATION NUMBER: US 08/542,363

PRIOR FILING DATE: 1995-10-12

PRIOR APPLICATION NUMBER: US 08/160,861

PRIOR FILING DATE: 1993-12-03

NUMBER OF SEQ ID NOS: 43

SOFTWARE: Patentin Version 3.0

SEQ ID NO: 4

LENGTH: 680

TYPE: PRT

ORGANISM: Homo sapiens

US-09-827-949-4

Query Match

Best Local Similarity 100.0%; Score 42; DB 4; Length 680;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKEV 9

Db 621 SLTANKREV 629

## RESULT 5

US-08-542-363-2  
Sequence 2, Application US/08542363

Patent No. 5770421

GENERAL INFORMATION:

APPLICANT: Morris, Stephan W.

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and

Ligands Thereof

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: STERN, KESSLER, GOLDSTEIN &amp; FOX P.L.L.C.

STREET: 1100 New York Avenue, N.W., Suite 600

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/542,363

FILING DATE: 12-OCT-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Fox, Samuel L.

REGISTRATION NUMBER: 30,353

REFERENCE/DOCKET NUMBER: 0656.0400001/SLF/GKT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1620 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-542-363-2

Query Match

Best Local Similarity 100.0%; Score 42; DB 1; Length 1620;

Matches 9; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

Db 1561 SLTANKREV 1569

RESULT 6

US-09-100-089-2

Sequence 2, Application US/09100089

Patent No. 6174674

GENERAL INFORMATION:

APPLICANT: Morris, Stephan W.

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and

Ligands Thereof

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: STERN, KESSLER, GOLDSTEIN &amp; FOX P.L.L.C.

STREET: 1100 New York Avenue, N.W., Suite 600

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/100,089

FILING DATE: Herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/160,861

FILING DATE: 03-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/542,363

FILING DATE: 12-OCT-1995

ATTORNEY/AGENT INFORMATION:

NAME: Fox, Samuel L.

REGISTRATION NUMBER: 30,353

REFERENCE/DOCKET NUMBER: 0656.0400002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1620 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-100-089-2

Query Match

Best Local Similarity 100.0%; Score 42; DB 3; Length 1620;

Matches 9; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

Db 1561 SLTANKREV 1569

RESULT 7

US-09-670-827-2

Sequence 2, Application US/09670827

Patent No. 645197

GENERAL INFORMATION:

APPLICANT: Morris, Stephan W.

TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and

Ligands Thereof

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: STERN, KESSLER, GOLDSTEIN &amp; FOX P.L.L.C.

STREET: 1100 New York Avenue, N.W., Suite 600

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/670,827

FILING DATE: 28-Sep-2000

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/160,861

FILING DATE: 03-DEC-1993

APPLICATION NUMBER: US 08/542,363

FILING DATE: 12-OCT-1995

APPLICATION NUMBER: US 09/100,089

FILING DATE: 19-JUN-1998

ATTORNEY/AGENT INFORMATION:

```
NAME: Fox, Samuel L.
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 0656.0400003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1620 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-670-827-2
Query Match 100.0%; Score 42; DB 4; Length 1620;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1 SLTANKEV 9
DB 1561 SLTANKEV 1569
RESULT 8
US-09-827-949-2
Sequence 2, Application US/09827949
Patent No. 6696548
GENERAL INFORMATION:
APPLICANT: Morris, Stephan W.
TITLE OF INVENTION: ALK Protein Tyrosine Kinase/Receptor and Ligands Thereof
FILE REFERENCE: 0656.0400004
CURRENT APPLICATION NUMBER: US/09/827,949
CURRENT FILING DATE: 2001-04-09
PRIOR APPLICATION NUMBER: US 09/670,827
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 09/100,089
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: US 08/542,363
PRIOR FILING DATE: 1995-10-12
PRIOR APPLICATION NUMBER: US 08/160,861
PRIOR FILING DATE: 1993-12-03
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 2
LENGTH: 1620
TYPE: PRT
ORGANISM: Homo sapiens
US-09-827-949-2
Query Match 100.0%; Score 42; DB 4; Length 1620;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1 SLTANKEV 9
DB 1561 SLTANKEV 1569
RESULT 9
US-09-134-000C-6610
Sequence 6610, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
```

```
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 6610
LENGTH: 378
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-134-000C-6610
Query Match 73.8%; Score 31; DB 4; Length 378;
Best Local Similarity 66.7%; Pred. No. 65;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
CY 1 SLTANKEV 9
DB 79 SLTANKEV 87
RESULT 10
US-09-710-279-1808
Sequence 1808, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMBERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: P03480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 1808
LENGTH: 179
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-1808
Query Match 71.4%; Score 30; DB 4; Length 179;
Best Local Similarity 75.0%; Pred. No. 46;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
CY 1 SLTANKEV 8
DB 51 SLTANKEV 58
RESULT 11
US-09-134-001C-4215
Sequence 4215, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO: 4215
LENGTH: 180
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4215
Query Match 71.4%; Score 30; DB 3; Length 180;
```

Best Local Similarity 75.0%; Pred. No. 47;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLTANKEV 8  
Db 52 SLTANKEV 59

RESULT 12  
US-09-583-110-3178  
; Sequence 3178, Application US/09583110  
; Patent No. 6699703  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al.  
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
; FILE REFERENCE: PATH00-07A  
; CURRENT APPLICATION NUMBER: US/09/583,110  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/107,433  
; PRIOR FILING DATE: 1998-06-30  
; PRIOR APPLICATION NUMBER: US 60/085,131  
; PRIOR FILING DATE: 1998-05-12  
; PRIOR APPLICATION NUMBER: US 60/051,553  
; PRIOR FILING DATE: 1997-07-02  
; NUMBER OF SEQ ID NOS: 5322  
; SEQ ID NO 3178  
; LENGTH: 232  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-583-110-3178

Query Match  
Best Local Similarity 71.4%; Score 30; DB 4; Length 232;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LTANKEV 9  
Db 86 LTANKEV 93

RESULT 13  
US-09-583-110-4472  
; Sequence 4472, Application US/09583110  
; Patent No. 6699703  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al.  
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
; FILE REFERENCE: PATH00-07A  
; CURRENT APPLICATION NUMBER: US/09/583,110  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/107,433  
; PRIOR FILING DATE: 1998-06-30  
; PRIOR APPLICATION NUMBER: US 60/085,131  
; PRIOR FILING DATE: 1998-05-12  
; PRIOR APPLICATION NUMBER: US 60/051,553  
; PRIOR FILING DATE: 1997-07-02  
; NUMBER OF SEQ ID NOS: 5322  
; SEQ ID NO 4472  
; LENGTH: 232  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-583-110-4472

Query Match  
Best Local Similarity 71.4%; Score 30; DB 4; Length 232;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LTANKEV 9  
Db 86 LTANKEV 93

RESULT 14  
US-09-252-991A-16966  
; Sequence 16966, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfeld et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 16966  
; LENGTH: 303  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-16966

Query Match  
Best Local Similarity 71.4%; Score 30; DB 4; Length 303;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
Db 64 SLTANKEV 72

RESULT 15  
US-09-710-279-320  
; Sequence 320, Application US/09710279  
; Patent No. 6703492  
; GENERAL INFORMATION:  
; APPLICANT: KIMBERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: P03460US  
; CURRENT APPLICATION NUMBER: US/09/710,279  
; PRIOR FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: Patent Ver. 2.1  
; SEQ ID NO 320  
; LENGTH: 338  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-710-279-320

Query Match  
Best Local Similarity 71.4%; Score 30; DB 4; Length 338;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
Db 329 SLTANKEV 337

Search completed: November 30, 2004, 07:19:17  
Job time : 23.6154 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 08:27:53 ; Search time 50 Seconds  
(without alignments)  
17.319 Million cell updates/sec

Title: US-10-008-377a-7  
Perfect score: 42  
Sequence: 1 SLTANKEV 9

Scoring table: BIOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283416 segs, 96216763 residues

Total number of hits satisfying chosen parameters: 791

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	45.2	8	1	IFSAME
2	16	39.1	6	2	cytotoxic T-lympho
3	15	35.7	5	2	hypoxanthine phosph
4	14	33.3	4	2	endoglucanase F
5	14	33.3	8	2	variant surface gl
6	13	31.0	8	2	gene Tinslow prote
7	13	31.0	8	2	glycine reductase
8	13	31.0	9	2	phosphoenolpyruvat
9	12	28.6	4	2	hypochlorite prote
10	12	28.6	5	2	lg heavy chain CRD
11	12	28.6	7	2	platelet glycoproc
12	12	28.6	7	2	amine oxidase (cop
13	12	28.6	7	2	Na+-transporting A
14	12	28.6	9	2	glutathione trans
15	11	26.2	4	2	hypochlorite prote
16	11	26.2	4	2	T-cell receptor be
17	11	26.2	5	2	T-cell receptor be
18	11	26.2	6	2	paraoxonol crystal
19	11	26.2	7	2	ribosomal protein
20	11	26.2	7	2	hypochlorite prote
21	11	26.2	7	2	phospholipase C (B
22	11	26.2	7	2	phospholipase C (B
23	11	26.2	7	2	phospholipase C (B
24	11	26.2	8	2	neuropeptide led-C
25	11	26.2	8	2	hypertrichosemic
26	11	26.2	8	2	adipokinetic hormo
27	11	26.2	8	2	P element, P cytot
28	11	26.2	8	2	cytochrome P450 AL
29	11	26.2	8	2	aspartate kinase (

30	11	26.2	8	2	SL9288	acylase - Kluver
31	11	26.2	9	2	PT0670	T-cell receptor be
32	11	26.2	9	2	B30572	T-cell receptor be
33	11	26.2	9	2	PH0902	T-cell receptor be
34	11	26.2	9	2	S78426	52.5K protein - sp
35	11	26.2	9	2	S56004	glucan 1,3-beta-gl
36	11	26.2	9	2	S19523	orf AB protein - S
37	11	26.2	9	2	PC7076	spectrin alpha cha
38	11	26.2	9	2	PC7078	unidentified 48.7K
39	11	26.2	9	2	B24362	chloramphenicol O-
40	11	26.2	9	2	S30494	cat gene leader pe
41	10	23.8	3	3	S13894	histidinol dehydro
42	10	23.8	3	3	PT0578	T-cell receptor be
43	10	23.8	4	2	140870	phospholipase C (B
44	10	23.8	4	2	S17255	ribosomal protein
45	10	23.8	5	2	B37325	pap fibribrial regul

## ALIGNMENTS

RESULT 1  
IFSAME  
probable msrA leader peptide - Staphylococcus epidermidis  
C:Species: Staphylococcus epidermidis  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C:Accession: S11157  
R:Rose, J.I.; Bady, E.A.; Cove, J.R.; Cunliffe, W.J.; Baumberg, S.; Mootron, J.C.  
Mol. Microbiol. 4, 1207-1214, 1990  
A>Title: Inducible erythromycin resistance in staphylococci is encoded by a member of the  
A:Reference number: S11157; MIMD:31041730; PMID:2233255  
A:Accession: S11157  
A:Molecule type: DNA  
A:Residues: 1-8 <ROS>  
A:Cross-references: UNIPROT:P23211; EMBL:X52085; NID:947000; PIDN:CA36303.1; PID:958165  
C:Superfamily: probable msrA leader peptide

Query Match 45.2% Score 19; DB 1; Length 8;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LTANMK 7  
DB 1 MTASMR 6

RESULT 2  
149424  
cytotoxic T-lymphocyte proteinase 3 (BC 3.4.21.-) - western wild mouse (fragment)  
C:Species: Mus spretus (western wild mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999  
C:Accession: I49424  
R:Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maezaki, Y.; Nadeau, J.F.  
Mamm. Genome 5, 349-355, 1994  
A>Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.  
A:Reference number: I48934; MIMD:94319082; PMID:8043949  
A:Accession: I49424  
A:Status: preliminary; translated from GB/EMBL/DDBT  
A:Molecule type: DNA  
A:Residues: 1-6 <RES>  
A:Cross-references: EMBL:U05745; NID:9497084; PIDN:AA60481.1; PID:9642831  
C:Keywords: hydrolase; serine proteinase

Query Match 38.1% Score 16; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 NMK 7  
DB 2 NMK 4

RESULT 3

A37114  
hypoxanthine phosphoribosyltransferase (EC 2.4.2.8) - fluke (Schistosoma mansoni) (fragm  
C/Species: Schistosoma mansoni  
C/Date: 28-Mar-1991 #sequence\_revision 28-Mar-1991 #text\_change 23-Jun-1993  
C/Accession: A37114  
R/Yuan, L.; Craig, S.P.; McKerrow, J.H.; Wang, C.C.  
J. Biol. Chem. 265, 13528-13532, 1990  
A/Title: The hypoxanthine-guanine phosphoribosyltransferase of Schistosoma mansoni. Funct  
A/Reference number: A37114; MUID:90337955; PMID:2199439  
A/Accession: A37114  
A/Status: Preliminary  
A/Molecule type: protein  
A/Residues: 1-5 <YDA>  
C/Keywords: glycosyltransferase; pentosyltransferase

Query Match 35.7%; Score 15; DB 2; Length 5;  
Best Local Similarity 40.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LITAM 6  
DB 1 MSSNM 5

RESULT 4  
endoglucanase F - Clostridium thermocellum (fragment)  
C/Species: Clostridium thermocellum  
C/Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 09-Jul-2004  
C/Accession: I40804  
R/Mishra, S.; Beguin, P.; Aubert, J.  
J. Bacteriol. 173, 80-85, 1991  
A/Title: Transcription of clostridium thermocellum endoglucanase genes celf and celd.  
A/Reference number: I40804; MUID:91100322; PMID:1987137  
A/Accession: I40804  
A/Status: Preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-4 <RES>  
A/Cross-references: UNIPROT:P26224; GB:M64363; NID:G144771  
C/Genetics:  
A/Genes: celf  
A/Start codon: TTG

Query Match 33.3%; Score 14; DB 2; Length 4;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 MKEV 9  
DB 1 MKKI 4

RESULT 5  
A21440  
variant surface glycoprotein pSLc1 - Trypanosoma brucei (fragment)  
C/Species: Trypanosoma brucei  
C/Date: 19-Nov-1988 #sequence\_revision 19-Nov-1988 #text\_change 09-Jul-2004  
C/Accession: A21440  
R/Barnoms, W.; Nelson, R.G.; Watkins, K.P.; Agabian, N.  
Cell 38, 309-316, 1984  
A/Title: Trypanosome mRNAs share a common 5' spliced leader sequence.  
A/Reference number: A90853; MUID:84282716; PMID:6088073  
A/Accession: A21440  
A/Molecule type: mRNA  
A/Residues: 1-8 <PAR>  
A/Cross-references: UNIPROT:P22225; GB:K02195; NID:9162150; PID:9162151  
C/Keywords: glycoprotein

Query Match 33.3%; Score 14; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 KEV 9

DB 4 KEV 6

RESULT 6  
gene Tnslow protein - rat (fragment)  
C/Species: Rattus sp. (rat)  
C/Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 28-Feb-1997  
C/Accession: I57532  
R/Banerjee-Basu, S.; Buonanno, A.  
Mol. Cell. Biol. 13, 7019-7028, 1993  
A/Title: cis-acting sequences of the rat tropomyosin I slow gene confer tissue- and develop  
A/Reference number: I57532; MUID:94019373; PMID:8413291  
A/Accession: I57532  
A/Status: Preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-8 <RES>  
A/Cross-references: GB:S66172; NID:G432603  
C/Genetics:  
A/Genes: Tnslow

Query Match 31.0%; Score 13; DB 2; Length 8;  
Best Local Similarity 75.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 MKEV 9  
DB 1 MPEV 4

RESULT 7  
A39308  
glycine reductase (EC 1.4.99.-) sulphydryl protein C, alpha chain - Clostridium sticklandii  
C/Species: Clostridium sticklandii  
C/Date: 19-Jun-1992 #sequence\_revision 19-Jun-1992 #text\_change 09-Jul-2004  
C/Accession: A39308  
R/Stadman, T.C.; Davis, J.N.  
J. Biol. Chem. 266, 22147-22153, 1991  
A/Title: Glycine reductase protein C. Properties and characterization of its role in the  
A/Reference number: A39308; MUID:92042141; PMID:1939235  
A/Accession: A39308  
A/Status: Preliminary  
A/Molecule type: protein  
A/Residues: 1-8 <STA>  
A/Cross-references: UNIPROT:Q7M0L0  
C/Function:  
A/Description: glycine reductase complex catalyzes the reductive deamination of glycine  
C/Keywords: ATP; oxidoreductase

Query Match 31.0%; Score 13; DB 2; Length 8;  
Best Local Similarity 75.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 MKEV 9  
DB 1 MKPV 4

RESULT 8  
S13889  
phosphoenolpyruvate carboxylase (EC 4.1.1.31) - maize  
C/Species: Zea mays (maize)  
C/Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C/Accession: S13889  
R/Jiao, Y.; Chollet, R.  
Arch. Biochem. Biophys. 283, 300-305, 1990  
A/Title: Regulatory phosphorylation of serine-15 in maize phosphoenolpyruvate carboxyla  
A/Reference number: S13889; MUID:91112741; PMID:2148563  
A/Accession: S13889  
A/Status: Preliminary  
A/Molecule type: protein  
A/Residues: 1-9 <JIA>

A:Cross-references: UNIPROT:Q43267; UNIPROT:Q41197  
 C:Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 31.0%; Score 13; DB 2; Length 9;  
 Best Local Similarity 28.6%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTANK 7  
 | : | : | :  
 Db 3 SIDAQLR 9

RESULT 9  
 S43014  
 hypothetical protein URF-2Y - Yersinia enterocolitica transposon TN3926  
 C:Species: Yersinia enterocolitica  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Oct-1999  
 C:Accession: S43014

R:Osbourn, S.E.V.; Turner, A.K.; Grinstead, J.  
 submitted to the EMBL Data Library, March 1994  
 A:Description: The structure of the bacterial transposable element, TN3926.  
 A:Reference number: S43011

A:Accession: S43014  
 A:Molecule type: DNA  
 A:Residues: 1-4 <OSB>  
 A:Cross-references: EMBL:X78059; NID:g460067; PIDN:CAAC4980.1; PID:g581836  
 C:Keywords:  
 A:Mobile element: transposon TN3926

Query Match 28.6%; Score 12; DB 2; Length 4;  
 Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LRAN 5  
 : | : | : |  
 Db 1 NNAN 4

RESULT 10  
 PT0295  
 Ig heavy chain CRD3 region (clone 5-91) - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
 C:Accession: PT0295

R:Yamada, M.; Messerman, R.; Reichard, B.A.; Shale, S.; Caton, A.J.; Rivera, G.  
 J. Exp. Med. 173:395-407, 1991  
 A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and J  
 A:Reference number: PT0222; MUID:91108337; PMID:1899102

A:Accession: PT0295  
 A:Molecule type: DNA  
 A:Residues: 1-5 <YAM>  
 A:Experimental source: B lymphocyte  
 C:Keywords: heterotrimer; immunoglobulin

Query Match 28.6%; Score 12; DB 2; Length 5;  
 Best Local Similarity 60.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTAN 5  
 | : | : | :  
 Db 1 SKTRN 5

RESULT 11

C56793  
 platelet glycoprotein GPIIa - human (fragment)

C:Species: Homo sapiens (man)  
 C:Date: 05-Jan-1996 #sequence\_revision 05-Jan-1996 #text\_change 07-Feb-1997  
 C:Accession: C56793

R:Catmell, B.; Parenttler, S.; Leung, L.L.; McGregor, J.L.  
 Biochem. J. 279, 419-425, 1991

A:Title: Separation of important new platelet glycoproteins (GPIIa, GPIc, GPIc\*, GPIIa and  
 A:Reference number: A56793; MUID:92061944; PMID:1953640

A:Accession: C56793  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-7 <CAT>  
 A:Experimental source: platelet  
 C:Keywords: glycoprotein

Query Match 28.6%; Score 12; DB 2; Length 7;  
 Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 NMKE 8  
 | : | : | :  
 Db 2 NLDE 5

RESULT 12  
 A38081  
 amine oxidase (copper-containing) (EC 1.4.3.6) - yeast (Pichia angusta) (fragment)

C:Species: Pichia angusta  
 C:Date: 31-Dec-1993 #sequence\_revision 03-Feb-1994 #text\_change 20-Apr-2000  
 C:Accession: A38081

R:Mu, D.; Jones, S.M.; Smith, A.J.; Brown, D.E.; Dooley, D.M.; Kilman, J.P.  
 J. Biol. Chem. 267, 7979-7982, 1992

A:Title: Tyrosine codon corresponds to topa quinone at the active site of copper amine c  
 A:Reference number: A38081; MUID:92235001; PMID:1569055

A:Accession: A38081  
 A:Molecule type: protein

A:Residues: 1-7 <MOA>  
 C:Keywords: copper; oxidoreductase; quinoprotein; topaquinone  
 F/4/Modified site: topaquinone (Tyr) #status experimental

Query Match 28.6%; Score 12; DB 2; Length 7;  
 Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 ANMKEV 9  
 | : | : | :  
 Db 2 ANYEV 7

RESULT 13  
 S45648

Nat-transporting ATP synthase (EC 3.6.1.-) alpha chain - Acetobacterium woodii (fragment)

N:Alternative names: ATPase alpha chain  
 C:Species: Acetobacterium woodii  
 C:Date: 19-Mar-1997 #sequence\_revision 06-Jun-1997 #text\_change 07-May-1999  
 C:Accession: S45648

R:Reidlinger, J.; Mueller, V.  
 Eur. J. Biochem. 223, 275-283, 1994

A:Title: Purification of ATP synthase from Acetobacterium woodii and identification as  
 A:Reference number: S45648; MUID:94307271; PMID:8033902

A:Accession: S45648  
 A:Molecule type: protein

A:Residues: 1-3/4-7 <REI>  
 A:Experimental source: DSM 1030

C:Keywords: hydrolase

Query Match 28.6%; Score 12; DB 2; Length 7;  
 Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 NMKE 8  
 | : | : | :  
 Db 2 NLDE 5

RESULT 14  
 S59902

glutathione transferase (EC 2.5.1.18) P - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 20-Jul-1996 #sequence\_revision 13-Mar-1997 #text\_change 17-Mar-1999  
 C:Accession: S59902

R:Nishihira, J.; Hibiy, Y.; Sakai, M.; Nishi, S.; Kumazaki, T.; Ohki, S.; Sakamoto, W.  
 Biochim. Biophys. Acta 1252: 233-238, 1995  
 A>Title: The C-terminal region, Arg(201)-Gln(209), of glutathione transferase P contribu  
 A:Reference number: S59902; MUID:96049505; PMID:7578228  
 A:Accession: S59902  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-9 <NIS>  
 C:Superfamily: glutathione transferase  
 C:Keywords: transferase

Query Match 28.6%; Score 12; DB 2; Length 9;  
 Best Local Similarity 28.6%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 LTANKE 8  
 : | :  
 DB 3 INNGKQ 9

## RESULT 15

T46627  
 hypothetical protein c4 - loblolly pine  
 C:Species: Pinus taeda (loblolly pine)  
 C>Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 18-Feb-2000  
 C:Accession: T46627  
 R:Chang, S.; Purysa, J.; Funkhouser, E.A.; Newton, R.J.; Cairney, J.  
 Submitted to the EMBL Data Library, July 1995  
 A:Description: Cloning of a chitinase homolog which lacks chitin binding sites and is d  
 A:Reference number: 223105  
 A:Accession: T46627  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-4 <CHA>  
 A:Cross-references: EMBL:U1309; NID:G974285; PID:G974282  
 A:Experimental source: Strain 66PT2x6PT3; 8 month seedlings

Query Match 26.2%; Score 11; DB 2; Length 4;  
 Best Local Similarity 75.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 MKEV 9  
 : | :  
 DB 1 MKLV 4

Search completed: November 30, 2004, 08:35:17  
 Job time : 51 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 08:16:03; Search time 58 Seconds  
(without alignments)  
89.282 Million cell updates/sec

Title: US-10-008-377a-7  
Perfect score: 42  
Sequence: 1 SLTANKKEV 9

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1598

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: UniProt\_02:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	45.2	8	1	LPMS STAP
2	13	45.2	8	2	Q7DKL7
3	13	45.2	8	2	BAA34538
4	17	40.5	8	2	Q9AGP4
5	16	38.1	9	2	Q78337
6	16	38.1	9	2	O81964
7	16	38.1	9	2	O81966
8	16	38.1	9	2	O81968
9	16	38.1	9	2	O81968
10	16	38.1	9	2	O81968
11	16	38.1	9	2	O81968
12	16	38.1	9	2	O81968
13	16	38.1	9	2	O81968
14	16	38.1	9	2	O81968
15	16	38.1	9	2	O81968
16	16	38.1	9	2	O81968
17	16	38.1	9	2	O81968
18	16	38.1	9	2	O81968
19	16	38.1	9	2	O81968
20	16	38.1	9	2	O81968
21	16	38.1	9	2	O81968
22	16	38.1	9	2	O81968
23	16	38.1	9	2	O81968
24	16	38.1	9	2	O81968
25	16	38.1	9	2	O81968
26	16	38.1	9	2	O81968
27	16	38.1	9	2	O81968
28	16	38.1	9	2	O81968
29	16	38.1	9	2	O81968
30	16	38.1	9	2	O81968
31	16	38.1	9	2	O81968

32	15	35.7	8	2	Q7GEP5	Q7GEP5 nicotiana b
33	15	35.7	8	2	Q36898	Q36898 nicotiana p
34	15	35.7	8	2	Q799V9	Q799V9 bacillus sp
35	15	35.7	9	2	O6LB30	O6LB30 homo sapien
36	15	35.7	9	2	O65710	O65710 rous sarcos
37	15	35.7	9	2	CA557977	CA557977 homo sapi
38	14	33.3	8	2	Q16428	Q16428 homo sapien
39	14	33.3	8	2	Q16468	Q16468 homo sapien
40	14	33.3	8	2	Q9N6M5	Q9N6M5 toxoplasma
41	14	33.3	8	2	Q7JK31	Q7JK31 gorilla gor
42	14	33.3	8	2	Q7JK32	Q7JK32 pan troglod
43	14	33.3	8	2	Q9MYL5	Q9MYL5 pongo pygma
44	14	33.3	8	2	P92211	P92211 agropyron c
45	14	33.3	8	2	P92215	P92215 amblyopyrum

## ALIGNMENTS

## RESULT 1

LPMS STAP STANDARD; PRT; 8 AA.

AC P23211;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 01-MAR-1992 (Rel. 21, Last annotation update)  
DE Probable msa leader peptide.  
OS Staphylococcus epidermidis.  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_Taxid=1282;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=968;  
RX MEDLINE=91041730; PubMed=2233255;  
RA Rose J.I., Eady E.A., Cove J.H., Cunliffe W.J., Baumberg S.,  
RW Woolton J.C.;  
RT "Inducible erythromycin resistance in staphylococci is encoded by a  
RT member of the ATP-binding transport super-gene family."  
RL Mol. Microbiol. 4:1207-1214(1990).  
CC -!- FUNCTION: May regulate expression of the erythromycin resistance  
CC protein.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: X52085; CAA36303.1; -  
CC DR PIR: S11157; LPSAMP.  
CC KW Leader peptide; Plasmid.  
CC SQ SEQUENCE 8 AA; 937 MW; FA37340685BDC1A6 CRC64;

Query Match 45.2%; Score 19; DB 1; Length 8;  
Best Local Similarity 50.0%; Pred. No. 1.8e+06;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LTANKK 7  
Db 1 MTSNR 6

RESULT 2

ID Q7DKL7 PRELIMINARY; PRT; 8 AA.

AC Q7DKL7;  
DT 05-JUL-2004 (TREMBLrel. 27, Created)  
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
DE Leader peptide.

OS Staphylococcus aureus.  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OK NCBI\_TaxID=1280;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MS968;  
 RX MEDLINE=99026951; PubMed=9809423;  
 RA Matsuo K., Endou K., Kobayashi H., Inoue M., Nakajima Y.;  
 RT "A plasmid that encodes three genes for resistance to macrolide  
 antibiotics in Staphylococcus aureus";  
 RL FEMS Microbiol. Lett. 167:221-227(1998).  
 DR EMBL, AB013298; BAA34538.1; FA37340685BDC1A6 CRC64;  
 SQ SEQUENCE 8 AA; 937 MW; FA37340685BDC1A6 CRC64;

Query Match 45.2%; Score 19; DB 2; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LTANMK 7  
 DB 1 MTASMR 6

RESULT 3  
 ID BAA34538 PRELIMINARY; PRT; 8 AA.  
 AC BAA34538;  
 DT 02-MAR-2004 (TREMBlrel. 27, Created)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)  
 DE Leader peptide.  
 OS Staphylococcus aureus.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Staphylococcus.  
 OK NCBI\_TaxID=1280;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MS968;  
 RA Matsuo K., Endou K., Kobayashi H., Inoue M., Nakajima Y.;  
 RT "A plasmid that encodes three genes for resistance to macrolide  
 antibiotics in Staphylococcus aureus";  
 RL FEMS Microbiol. Lett. 167:221-227(1998).  
 DR EMBL, AB013298; BAA34538.1;  
 SQ SEQUENCE 8 AA; 937 MW; FA37340685BDC1A6 CRC64;

Query Match 45.2%; Score 19; DB 2; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LTANMK 7  
 DB 1 MTASMR 6

RESULT 4  
 ID Q9AGP4 PRELIMINARY; PRT; 8 AA.  
 AC Q9AGP4;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DE Serine hydroxymethyltransferase (Fragment).  
 GN Name=g1yA;  
 OS Arabidopsis sp. 11N.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Micrococcales; Micrococcaceae; Arthrobacter.  
 OK NCBI\_TaxID=153502;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=11N;  
 RA Meskys R., Harris R.J., Casalte V., Baerian J., Scrutton N.S.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, AF329478; AAK16486.1; -.

DR GO:0008168; F-methyltransferase activity; IEA.  
 DR GO:0016740; F:transferase activity; IEA.  
 KM Methyltransferase; Transferase.  
 FT NON\_TER  
 SQ SEQUENCE 8 AA; 898 MW; 6B1870533372457 CRC64;

Query Match 40.5%; Score 17; DB 2; Length 8;  
 Best Local Similarity 40.0%; Pred. No. 1.8e+06;  
 Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 NMKEV 9  
 DB 2 NIKKI 6

RESULT 5  
 ID Q78337 PRELIMINARY; PRT; 9 AA.  
 AC Q78337; 082778;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Ribulose 1,5-bisphosphate carboxylase/oxygenase large subunit (RbcL)  
 DE (Fragment).  
 GN Name=rdcl;  
 OS Caloglossa lepreurii.  
 OC Chlorophyta.  
 OC Eukaryota; Rhodophyta; Florideophyceae; Ceramiales; Delesseriaceae;  
 OC Caloglossa.  
 OK NCBI\_TaxID=76903;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=736, 1048, 902, 490, 932, 922, 9, 880, 1053, and 1052;  
 RA Kamiya M., West J.A., King R.J., Zuccarello G.C., Tanaka U., Hara Y.;  
 RT "Evolutionary divergence in the red algae Caloglossa lepreurii and C.  
 apometotica";  
 RL J. Phycol. 34:361-370(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=736, 1048, 902, 490, 932, 922, 9, 880, 1053, and 1052;  
 RA Kamiya M., Tanaka U., King R.J., West J.A., Zuccarello G.C., Kawai H.;  
 RT "Reproductive and genetic distinction between broad and narrow  
 ecotypes of Caloglossa continua (Delesseriaceae, Rhodophyta).";  
 RL Phycologia 38:356-367(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Virginia USA, and Georgia USA;  
 RA Kamiya M., West J.A., Zuccarello G.C., Kawai H.;  
 RT "Caloglossa intermedia, sp. nov. (Rhodophyta) from the western  
 Atlantic coast: molecular and morphological analyses with special  
 reference to C. lepreurii and C. monosticha.";  
 RL J. Phycol. 36:411-420(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Zuccarello G.C., Bartlett J., Yeates P.H.;  
 RT "Differentiation of Caloglossa lepreurii (Rhodophyta) populations in  
 Bur. J. Phycol. 0:0-0(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA West J.A., Zuccarello G.C., Kamiya M.;  
 RT "Reproductive patterns of Caloglossa species (Delesseriaceae,  
 Rhodophyta) from Australia and New Zealand: multiple origins of  
 asexuality in C. lepreurii. Literature review on aponmixis, mixed-  
 phase, bisexuality and sexual compatibility";  
 RL Phycol. Res. 49:183-200(2001).  
 DR EMBL, D87813; BAA31279.1; -  
 DR EMBL, D89959; BAA31303.1; -  
 DR EMBL, AB030267; BAA90255.1; -  
 DR EMBL, AB030268; BAA90257.1; -  
 DR EMBL, D89949; BAA31283.1; -  
 DR EMBL, D89951; BAA31287.1; -  
 DR EMBL, D89952; BAA31289.1; -.

DR EMBL: D89953; BAA31291.1; -  
 DR EMBL: D89954; BAA31293.1; -  
 DR EMBL: D89955; BAA31295.1; -  
 DR EMBL: D89956; BAA31297.1; -  
 DR EMBL: D89957; BAA31299.1; -  
 DR EMBL: D89958; BAA31301.1; -  
 DR EMBL: AF242423; AAG35289.1; -  
 DR EMBL: AF242424; AAG35291.1; -  
 DR EMBL: AF242425; AAG35293.1; -  
 DR EMBL: AF242426; AAG35295.1; -  
 DR EMBL: AF242427; AAG35297.1; -  
 DR EMBL: AF242428; AAG35299.1; -  
 DR EMBL: AF242429; AAG35301.1; -  
 DR EMBL: AF340169; AAK98066.1; -  
 DR EMBL: AF340170; AAK98067.1; -  
 DR EMBL: AF340171; AAK98069.1; -  
 DR EMBL: AF340172; AAK98071.1; -  
 DR EMBL: AF340173; AAK98073.1; -  
 DR EMBL: AF340174; AAK98075.1; -  
 DR EMBL: AF340175; AAK98077.1; -  
 DR EMBL: AF340176; AAK98079.1; -  
 DR EMBL: AF340177; AAK98081.1; -  
 DR GO: 0009507; C:chloroplast; IEA.  
 KM Chloroplast.  
 FT NON TER 1  
 SQ SEQUENCE 9 AA; 977 MW; CA1A4DC1B771AB02 CRC64;

Query Match 38.1%; Score 16; DB 2; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

CY 3 TANM 6  
 |||:  
 6 TANV 9

RESULT 6  
 ID 081964 PRELIMINARY; PRT; 9 AA.  
 AC 081964; Q9T1381;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Ribulose 1,5-bisphosphate carboxylase/oxygenase large subunit  
 DE (Fragment).  
 GN Name: rbcL;  
 OS Caloglossa continua.  
 OC Caloglossa.  
 OC Eukaryota; Rhodophyta; Florideophyceae; Ceramiales; Delesseriaceae;  
 OC Caloglossa.  
 OX NCBI\_TaxID=76905;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=639;  
 RA Kamiya M., West J.A., King R.J., Zuccarello G.C., Tanaka J., Hara Y.,  
 RT "Evolutionary divergence in the red algae Caloglossa lepteurii and C.  
 RT J. Physiol. 34:361-370(1998).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=639, 500, and 729;  
 RA Kamiya M., Tanaka J., King R.J., West J.A., Zuccarello G.C., Kawai H.;  
 RT "Reproductive and genetic distinction between broad and narrow  
 RT entities of Caloglossa continua (Delesseriaceae, Rhodophyta).";  
 RL Phycologia 38:356-367(1999).  
 DR EMBL: D89950; BAA31285.1; -  
 DR EMBL: AB023379; BAA8908.1; -  
 DR EMBL: AB023380; BAA8910.1; -  
 DR GO: 0009507; C:chloroplast; IEA.  
 KM Chloroplast.  
 FT NON TER 1  
 SQ SEQUENCE 9 AA; 977 MW; CA1A4DC1B771AB02 CRC64;

Query Match 38.1%; Score 16; DB 2; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

CY 3 TANM 6  
 |||:  
 6 TANV 9

RESULT 7

ID 081966 PRELIMINARY; PRT; 9 AA.  
 AC 081966; Q9T1381;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Ribulose 1,5-bisphosphate carboxylase/oxygenase large subunit  
 DE (Fragment).  
 GN Name: rbcL;  
 OS Caloglossa monosticha.  
 OC Caloglossa.  
 OC Eukaryota; Rhodophyta; Florideophyceae; Ceramiales; Delesseriaceae;  
 OC Caloglossa.  
 OX NCBI\_TaxID=76906;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=892;  
 RA Kamiya M., West J.A., King R.J., Zuccarello G.C., Tanaka J., Hara Y.,  
 RT "Evolutionary divergence in the red algae Caloglossa lepteurii and C.  
 RT J. Physiol. 34:361-370(1998).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=892, and 987;  
 RA Kamiya M., Tanaka J., King R.J., West J.A., Zuccarello G.C., Kawai H.;  
 RT "Reproductive and genetic distinction between broad and narrow  
 RT entities of Caloglossa continua (Delesseriaceae, Rhodophyta).";  
 RL Phycologia 38:356-367(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D571;  
 RA West J.A., Zuccarello G.C., Kamiya M.;  
 RT "Reproductive patterns of Caloglossa species (Delesseriaceae,  
 RT Rhodophyta) from Australia and New Zealand: multiple origins of  
 RT asexuality in C. lepteurii. Literature review on apomixis, mixed-  
 RT phase, bisexuality and sexual compatibility.";  
 RL Physiol. Res. 49:183-200(2001).  
 DR EMBL: D89960; BAA31305.1; -  
 DR EMBL: AF340178; AAK98083.1; -  
 DR EMBL: AF340179; AAK98085.1; -  
 DR EMBL: AF340180; AAK98087.1; -  
 DR EMBL: AB023381; BAA8912.1; -  
 DR GO: 0009507; C:chloroplast; IEA.  
 KM Chloroplast.  
 FT NON TER 1  
 SQ SEQUENCE 9 AA; 977 MW; CA1A4DC1B771AB02 CRC64;

Query Match 38.1%; Score 16; DB 2; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

CY 3 TANM 6  
 |||:  
 6 TANV 9

RESULT 8

ID 081968 PRELIMINARY; PRT; 9 AA.  
 AC 081968; Q95B33;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

```

DE Ribulose 1,5-bisphosphate carboxylase/oxygenase large subunit
DE (Fragment).
GN Name=tbl;
OS Caloglossa ogasawarensis.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Florideophyceae; Ceramiales; Delesseriaceae;
OC Caloglossa.
OX NCBI_TaxID=76907;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=596;
RA Kamiya M., West J.A., King R.J., Zuccarello G.C., Tanaka J., Hara Y.;
RT "Reproductive patterns of Caloglossa species (Delesseriaceae,
RT Rhodophyta) from Australia and New Zealand: multiple origins of
RT asexuals in C. lepreurii. Literature review on apomixis, mixed-
RT phase, bisexuality and sexual compatibility.";
RL Phycol. Res. 49:183-200(2001).
DR EMBL; D89961; BA31307.1; -.
DR EMBL; AF340182; AA938091.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
KW Chloroplast.
FT NON TER
SQ SEQUENCE 9 AA; 977 MW; CA1A4DC1B771AB02 CRC64;

Query Match 38.1%; Score 16; DB 2; Length 9;
Best Local Similarity 75.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 3 TANM 6
DB 6 TANV 9

```

```

RESULT 9
ID Q76FS7 PRELIMINARY; PRT; 9 AA.
AC Q76FS7;
DT 05-JUL-2004 (TRENBLREL. 27, Created)
DT 05-JUL-2004 (TRENBLREL. 27, Last sequence update)
DE Ribulose-1,5-bisphosphate large subunit (Fragment).
GN Name=tbl;
OS Hypnea japonica.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Florideophyceae; Gigartinales; Hypneaceae;
OC Hypnea.
OX NCBI_TaxID=105606;
RN [1]
RP SEQUENCE FROM N.A.
RA Yano T., Kamiya M., Arai S., Kawai H.;
RT Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB104706; BA082415.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
KW Chloroplast.
FT NON TER
SQ SEQUENCE 9 AA; 977 MW; CA1A4DC1B771AB02 CRC64;

Query Match 38.1%; Score 16; DB 2; Length 9;
Best Local Similarity 75.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 3 TANM 6
DB 6 TANV 9

```

```

QY 3 TANM 6
DB 6 TANV 9

RESULT 10
ID Q8HB43 PRELIMINARY; PRT; 9 AA.
AC Q8HB43;
DT 01-MAR-2003 (TRENBLREL. 23, Created)
DT 01-MAR-2003 (TRENBLREL. 23, Last sequence update)
DE Ribulose-1,5-bisphosphate large subunit (Fragment).
GN Name=tbl;
OS Gracilaria perleixa.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Florideophyceae; Gracilariiales; Gracilariaceae;
OC Gracilaria.
OX NCBI_TaxID=197869;
RN [1]
RP SEQUENCE FROM N.A.
RA Byrne K., Zuccarello G.C., West J.A., Liao M.-L., Kraft G.;
RT "Gracilaria species (Gracilariaceae, Rhodophyta) from southeastern
RT Australia, including a new species, G. perleixa sp. nov.: Morphology,
RT molecular relationships and agar content.";
RL Phycol. Res. 50:295-311(2002).
DR EMBL; AY131306; AA07038.1; -.
DR EMBL; AY131307; AA07040.1; -.
DR EMBL; AY131308; AA07042.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
KW Chloroplast.
FT NON TER
SQ SEQUENCE 9 AA; 977 MW; CA1A4DC1B771AB02 CRC64;

Query Match 38.1%; Score 16; DB 2; Length 9;
Best Local Similarity 75.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 3 TANM 6
DB 6 TANV 9

```

```

RESULT 11
ID Q8HB44 PRELIMINARY; PRT; 9 AA.
AC Q8HB44;
DT 01-MAR-2003 (TRENBLREL. 23, Created)
DT 01-MAR-2003 (TRENBLREL. 23, Last sequence update)
DE Ribulose-1,5-bisphosphate large subunit (Fragment).
GN Name=tbl;
OS Gracilaria cliftonii.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Florideophyceae; Gracilariiales; Gracilariaceae;
OC Gracilaria.
OX NCBI_TaxID=206548;
RN [1]
RP SEQUENCE FROM N.A.
RA Byrne K., Zuccarello G.C., West J.A., Liao M.-L., Kraft G.;
RT "Gracilaria species (Gracilariaceae, Rhodophyta) from southeastern
RT Australia, including a new species, G. perleixa sp. nov.: Morphology,
RT molecular relationships and agar content.";
RL Phycol. Res. 50:295-311(2002).
DR EMBL; AY131302; AA07030.1; -.
DR EMBL; AY131303; AA07032.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
KW Chloroplast.
FT NON TER
SQ SEQUENCE 9 AA; 977 MW; CA1A4DC1B771AB02 CRC64;

Query Match 38.1%; Score 16; DB 2; Length 9;

```

Best Local Similarity 75.0%; Pred. No. 1.8e-06;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 TANM 6  
DB 6 TANV 9

## RESULT 12

OSHRJ8 PRELIMINARY; PRT; 9 AA.

AC OSHR45  
DT 01-MAR-2003 (TREMURel. 23, Created)  
DT 01-MAR-2003 (TREMURel. 23, Last sequence update)  
DT 05-JUN-2004 (TREMURel. 27, Last annotation update)  
DE Ribulose-1,5-bisphosphate large subunit (Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit) (Fragment).  
GN Name=rbcl;  
OS Gracilaria chilensis (Red alga).  
OC Chloroplast.  
CC Eukaryota; Rhodophyta; Florideophyceae; Gracilariiales; Gracilariaceae;  
CX NCBI\_TaxID=2775;  
RN 11

SEQUENCE FROM N.A.  
RA Byrne K., Zuccarello G.C., West J.A., Liao M.-L., Kraft G.,  
RT "Gracilaria species (Gracilariaceae, Rhodophyta) from southeastern Australia, including a new species, G. perplexa sp. nov.: Morphology, molecular relationships and agar content."  
RL Phycol. Res. 50:295-311(2002).  
RN 12

SEQUENCE FROM N.A.  
RA Cohen S., Fageron S., Martinez E., Correa J.A., Viard F.,  
RL Deetombe C., Valero M.;  
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY131299; AN07024.1; -  
DR EMBL: AY131300; AN07026.1; -  
DR EMBL: AY131301; AN07028.1; -  
DR EMBL: AY423840; AA84581.1; -  
DR GO:GO:0009507; C:chloroplast; IEA.  
KM Chloroplast.  
FT NON\_TER  
SQ SEQUENCE 9 AA; 977 MW; CA1A4DC1B771AB02 CRC64;

Query Match 38.1%; Score 16; DB 2; Length 9;  
Best Local Similarity 75.0%; Pred. No. 1.8e-06;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 TANM 6  
DB 6 TANV 9

## RESULT 13

OSHRJ8 PRELIMINARY; PRT; 9 AA.

AC OSHR45  
DT 01-MAR-2003 (TREMURel. 23, Created)  
DT 01-MAR-2003 (TREMURel. 23, Last sequence update)  
DT 01-JUN-2003 (TREMURel. 24, Last annotation update)  
DE Ribulose-1,5-bisphosphate large subunit (Fragment).  
GN Name=rbcl;  
OS Ptilophora prolifera.  
OC Chloroplast.  
CC Eukaryota; Rhodophyta; Florideophyceae; Gelidiales; Gelidiaceae;  
CX NCBI\_TaxID=143014;  
RN 11

SEQUENCE FROM N.A.  
RA Byrne K., Zuccarello G.C., West J.A., Liao M.-L., Kraft G.,  
RT "Gracilaria species (Gracilariaceae, Rhodophyta) from southeastern Australia, including a new species, G. perplexa sp. nov.: Morphology, molecular relationships and agar content."  
RN 12

RL Phycol. Res. 50:295-311(2002).

DR EMBL: AY131312; AN07050.1; -  
DR GO:GO:0009507; C:chloroplast; IEA.  
KM Chloroplast.  
FT NON\_TER  
SQ SEQUENCE 9 AA; 977 MW; CA1A4DC1B771AB02 CRC64;

Query Match 38.1%; Score 16; DB 2; Length 9;  
Best Local Similarity 75.0%; Pred. No. 1.8e-06;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 TANM 6  
DB 6 TANV 9

## RESULT 14

OSHRK0 PRELIMINARY; PRT; 9 AA.

AC OSHR45  
DT 01-MAR-2003 (TREMURel. 23, Created)  
DT 01-MAR-2003 (TREMURel. 23, Last sequence update)  
DT 01-JUN-2003 (TREMURel. 24, Last annotation update)  
DE Ribulose-1,5-bisphosphate large subunit (Fragment).  
GN Name=rbcl;  
OS Gracilaria secundata.  
OC Chloroplast.  
CC Eukaryota; Rhodophyta; Florideophyceae; Gracilariiales; Gracilariaceae;  
CX NCBI\_TaxID=172973;  
RN 11

SEQUENCE FROM N.A.  
RA Byrne K., Zuccarello G.C., West J.A., Liao M.-L., Kraft G.,  
RT "Gracilaria species (Gracilariaceae, Rhodophyta) from southeastern Australia, including a new species, G. perplexa sp. nov.: Morphology, molecular relationships and agar content."  
RL Phycol. Res. 50:295-311(2002).  
DR EMBL: AY131311; AN07048.1; -  
DR GO:GO:0009507; C:chloroplast; IEA.  
KM Chloroplast.  
FT NON\_TER  
SQ SEQUENCE 9 AA; 977 MW; CA1A4DC1B771AB02 CRC64;

Query Match 38.1%; Score 16; DB 2; Length 9;  
Best Local Similarity 75.0%; Pred. No. 1.8e-06;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 TANM 6  
DB 6 TANV 9

## RESULT 15

OSHSLO PRELIMINARY; PRT; 9 AA.

AC OSHSLO  
DT 01-MAR-2003 (TREMURel. 23, Created)  
DT 01-MAR-2003 (TREMURel. 23, Last sequence update)  
DT 01-JUN-2003 (TREMURel. 24, Last annotation update)  
DE Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (Fragment).  
GN Name=rbcl;  
OS Murayella pericladus.  
OC Chloroplast.  
CC Eukaryota; Rhodophyta; Florideophyceae; Ceramiales; Rhodometaceae;  
CX NCBI\_TaxID=110473;  
RN 11

SEQUENCE FROM N.A.  
RA Zuccarello G.C., Sanderoock B., West J.A.,  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF458764; AA014538.1; -  
DR GO:GO:0009507; C:chloroplast; IEA.

KW Chloroplast. 1  
 FT NON TER 1  
 SQ SEQUENCE 9 AA; 977 MW; CAA4ADC1B71AB02 CRC64;

Query Match 38.1%; Score 16; DB 2; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 TANN 6  
 |||:  
 Db 6 TANN 9

Search completed: November 30, 2004, 08:34:16  
 Job time : 60 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on: November 30, 2004, 08:14:33 ; Search time 114 Seconds  
(without alignments)  
28.321 Million cell updates/sec

Title: US-10-008-377A-7

Perfect score: 42

Sequence: 1 SLTANKKEV 9

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 313949

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*  
1: geneseqp1980s:\*  
2: geneseqp1980s:\*  
3: geneseqp2000s:\*  
4: geneseqp2000s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	42	100.0	9	AAO22989 Human p62
2	24	57.1	9	AAE07594 Human PUM
3	24	57.1	9	ABR23168 Human can
4	24	57.1	9	ABR21768 Human can
5	24	57.1	9	ABR20369 Human can
6	23	54.8	9	ABJ19931 MHC bindi
7	23	54.8	9	ABJ20118 MHC bindi
8	23	54.8	9	ABR23049 Human can
9	23	54.8	9	ABR20246 Human can
10	23	54.8	9	ABR21637 Human can
11	22	52.4	9	ADH68371 RSV F-pro
12	22	52.4	9	AAA66277 HPV type
13	22	52.4	9	ADH40759 Human CD4
14	22	52.4	9	ADH40364 Human CD4
15	21	50.0	9	AAE0252 HOP prote
16	21	50.0	9	ABM66287 Prolonib
17	20	47.6	9	ADH68370 RSV F-pro
18	20	47.6	7	AAE81479 Murine me
19	20	47.6	8	AAE98526 Human pep
20	20	47.6	7	AAE21665 Immunomod
21	20	47.6	8	AAE09862 Immunomod
22	20	47.6	8	AAE20699 Immunomod
23	20	47.6	8	ABR83118 Beta-gala
24	20	47.6	9	AAE57750 Peptide (
25	20	47.6	9	AAE55417 HLA bindi

26	20	47.6	9	AAE09858 Immunomod
27	20	47.6	9	AAE75636 HLA class
28	20	47.6	9	AAE20695 Immunomod
29	20	47.6	9	AAU94681 Human nov
30	20	47.6	9	AAU94477 Human nov
31	20	47.6	9	AAU94091 Human nov
32	20	47.6	9	ABR22026 Human can
33	20	47.6	9	ABR23439 Human can
34	20	47.6	9	ABR23216 Human can
35	20	47.6	9	ABR20255 Human can
36	20	47.6	9	ABR21647 Human can
37	20	47.6	9	ABR21809 Human can
38	20	47.6	9	ABR20636 Human can
39	20	47.6	9	ABR23063 Human can
40	20	47.6	9	ABR20415 Human can
41	20	47.6	9	ADK10387 Human pap
42	19	45.2	6	AAW83819 Peptide s
43	19	45.2	7	AAW83709 Peptide s
44	19	45.2	7	ABP49017 Zinc fing
45	19	45.2	7	ABP70385 Tryptic p

## ALIGNMENTS

RESULT 1  
AAO22989  
ID AAO22989 standard; peptide; 9 AA.  
XX  
AC AAO22989;  
XX  
DT 17-SEP-2003 (first entry)  
XX  
DE Human p621-629 ALK-derived HLA-A\*0201 restricted CTL epitope peptide.  
XX  
KW HLA-A\*0201 binding; anaplastic lymphoma kinase; ALK; CTL epitope; ALCL;  
XX cytotoxic T-lymphocyte; receptor tyrosine; NPM/ALK fusion protein; human;  
XX oncogenic; anaplastic large cell lymphoma; neuroblastoma; gene therapy;  
XX t(2;5) translocation; neoplasia; cytostatic; immunotherapy; alkoma;  
XX p621-629.  
XX  
CS Homo sapiens.  
XX  
EN WO2003042243-A2.  
XX  
PD 22-MAY-2003.  
XX  
PF 14-NOV-2002; 2002WO-EP012764.  
XX  
PR 15-NOV-2001; 2001US-00008377.  
XX  
PA (NAST-) 1ST NAZ STUDIO E CURA DEI TUMORI.  
XX (SUDD-) ST JUDE CHILDREN'S RES HOSPITAL.  
XX  
PI Gambacorti-Passerini C, Passoni L;  
XX  
DR WPI; 2003-441791/41.  
XX  
PT New HLA-A\*0201-binding Anaplastic Lymphoma Kinase (ALK) peptide,  
XX useful for preparing a composition for treating ALK-positive lymphoma,  
XX neuroblastoma or ALK-expressing neoplasia.  
XX  
PS Claim 1; Page 7; 33pp; English.  
XX  
CC The invention relates to a novel HLA-A\*0201-binding anaplastic lymphoma  
XX kinase (ALK)-derived peptide, of 9-10 amino acids, which may act as a  
XX cytotoxic T-lymphocyte (CTL)-epitope. ALK is a receptor tyrosine kinase  
XX which forms part of the NPM/ALK fusion protein (oncogenic lymphoma kinase  
XX fusion protein) in ALCL (anaplastic large cell lymphoma)-derived cell  
XX lines. More than 50% of ALCL cases possess a t(2;5) chromosomal  
XX translocation that leads to the expression of the NPM/ALK fusion protein  
XX which forms a potent oncogene when constitutively activated. Translocated  
XX ALK is a widely expressed tumour-associated antigen characteristic of ALK

CC -positive lymphomas, neuroblastomas and ALK-expressing neoplasias. The  
 CC peptides of the invention demonstrate cytostatic activity and induce an  
 CC MHC (major histocompatibility complex) class I restricted cytotoxic  
 CC lymphocyte response against tumour cells expressing the NPM/ALK fusion  
 CC protein. Hence, the peptides may be utilised during the treatment, via  
 CC immunotherapy, of ALK-positive lymphomas such as alkomas, neuroblastomas  
 CC and ALK-expressing neoplasias. In addition, the peptides may be used  
 CC during gene therapy. The current sequence is that of the human p621-629  
 CC ALK-derived HLA-A\*0201 restricted CTL epitope peptide of the invention  
 XX

SO Sequence 9 AA;

Query Match 100.0%; Score 42; DB 6; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLTANKEV 9  
 |||||  
 Db 1 SLTANKEV 9

RESULT 2  
 AAE07594  
 ID AAE07594 standard; peptide; 9 AA.  
 XX  
 AC AAE07594;  
 XX  
 D7 06-NOV-2001 (first entry)  
 XX

DE Human PUMP-1 peptide (residues 54-62).  
 XX

KM PUMP-1 protease; cancer; neoplastic state; malignancy; ovary; lung;  
 KM prostate; colon; cytostatic; gene therapy; vaccine; immunogen; human.  
 XX

OS Homo sapiens.  
 XX

PN WO200154712-A1.  
 XX

PD 02-AUG-2001.  
 XX

PF 26-JAN-2001; 2001WO-US002698.  
 XX

PR 27-JAN-2000; 2000US-00492543.  
 XX

PA (UYAR-) UNIV ARKANSAS.  
 XX

PI O'Brien TJ;  
 XX

DR WPI; 2001-496835/54.  
 XX

PT Diagnosing cancer in an individual, useful for early detection of ovarian  
 PT cancer, lung cancer or prostate cancer, comprises determining the  
 PT presence of PUMP-1 protease in biological sample obtained from  
 PT individual.  
 XX

PS Example 20; Page 65; 145pp; English.  
 XX

CC The patent discloses compositions and methods for the early diagnosis of  
 CC ovarian cancer. The method involves obtaining a biological sample from an  
 CC individual and detecting PUMP-1 protease in the sample, where the  
 CC presence of PUMP-1 in the sample is indicative of the presence of cancer  
 CC in the individual and the absence of PUMP-1 is indicative of the absence  
 CC of cancer in the individual. The method is useful for the early detection  
 CC or diagnosis of ovarian cancer and other neoplastic state or malignancies  
 CC (e.g. lung cancer, prostate cancer, colon cancer or other cancers in  
 CC which PUMP-1 is overexpressed). It is also useful for diagnosing whether  
 CC an individual has cancer, is suspected of having cancer or is at risk of  
 CC getting cancer. The method is also used to detect and treat malignant  
 CC hyperplasia. The PUMP-1 proteins are also useful for vaccinating against  
 CC neoplastic states. The oligonucleotide is useful for treating neoplastic  
 CC states such as lung cancer, prostate cancer, colon cancer or other  
 CC cancers in which PUMP-1 is overexpressed. The present sequence is PUMP-1  
 CC peptide (residues 54-62) from human. This peptide is a putative immunogen

CC and binds strongly to HLA A1 and is used to inoculate an individual  
 CC against PUMP-1

XX Sequence 9 AA;

SO Query Match 57.1%; Score 24; DB 4; Length 9;  
 Best Local Similarity 55.6%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SLTANKEV 9  
 |||||  
 Db 1 SLTANKEV 9

RESULT 3  
 ABR23168  
 ID ABR23168 standard; peptide; 9 AA.  
 XX  
 AC ABR23168;  
 XX

D7 19-MAY-2003 (first entry)  
 XX

DE Human cancer-related protein 185P2C9 HLA peptide #3003.  
 XX

KM Human; cytostatic; vaccine; cancer; immune response; HLA;  
 KM human leukocyte antigen.  
 XX

OS Homo sapiens.  
 XX

PN WO200283921-A2.  
 XX

PD 24-OCT-2002.  
 XX

PF 10-APR-2002; 2002WO-US011654.  
 XX

PR 10-APR-2001; 2001US-0282739P.  
 XX

PR 10-APR-2001; 2001US-0283112P.  
 XX

PA (AGEN-) AGENSYS INC.  
 XX

PI Jakobovits A, Challita-Bid PM, Faris M, Ge W, Hubert RS;  
 PI Morrison K, Morrison RK, Raitano AB;  
 XX

DR WPI; 2003-075555/07.  
 XX

PT New composition comprising a substance that modulates the structure of  
 PT proteins and polynucleotides, useful for therapeutic, prognostic and  
 PT diagnostic reagents for eliciting cellular or humoral immune response in  
 PT cancer patients.  
 XX

PS Claim 13; Page 354; 1021pp; English.  
 XX

CC The present invention relates to novel human cancer-related genes and  
 CC proteins (AB278120-AB278168 and ABR01789-ABR01861). The genes and  
 CC proteins are useful for eliciting a humoral or cellular immune response.  
 CC The genes are useful as probes and primers for the amplification and/or  
 CC detection of genes, mRNAs or their fragments, as reagents for the  
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of  
 CC directing the expression of the protein, as tools for modulating or  
 CC inhibiting the expression of genes and/or translation of transcripts, and  
 CC as therapeutic agents. The proteins and peptides are useful as  
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present  
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example  
 CC from the invention  
 XX

SO Sequence 9 AA;

Query Match 57.1%; Score 24; DB 6; Length 9;  
 Best Local Similarity 55.6%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SLTANKEV 9

Db 1 NLSDMKREV 9

RESULT 4  
ID ABR21768 standard; peptide; 9 AA.  
XX ABR21768;  
AC ABR21768;  
XX 19-MAY-2003 (first entry)  
XX  
XX  
XX Human cancer-related protein 185P2C9 HLA peptide #1603.  
XX  
XX Human; cytostatic; vaccine; cancer; immune response; HLA;  
XX human leukocyte antigen.  
XX Homo sapiens.  
XX WO200283921-A2.  
XX  
XX 24-OCT-2002.  
XX  
XX 10-APR-2002; 2002WO-US011654.  
XX  
XX 10-APR-2001; 2001US-0282739P.  
XX 10-APR-2001; 2001US-0283112P.  
XX 25-APR-2001; 2001US-0286630P.  
XX  
XX (AGEN-) AGENSYS INC.  
XX  
XX Jakobovits A, Chailita-Bid PM, Paris M, Ge W, Hubert RS;  
XX Morrison K, Morrison RK, Raitano AB;  
XX WPI; 2003-075555/07.  
XX  
XX New composition comprising a substance that modulates the structure of  
XX proteins and polynucleotides, useful for therapeutic, prognostic and  
XX diagnostic reagents for eliciting cellular or humoral immune response in  
XX cancer patients.  
XX  
XX Claim 13; Page 338; 1021pp; English.  
XX  
XX The present invention relates to novel human cancer-related genes and  
XX proteins (AB278120-AB278168 and ABR01789-ABR01861). The genes and  
XX proteins are useful for eliciting a humoral or cellular immune response.  
XX The genes are useful as probes and primers for the amplification and/or  
XX detection of genes, mRNAs or their fragments, as reagents for the  
XX diagnosis and/or prognosis of cancer, as coding sequences capable of  
XX directing the expression of the protein, as tools for modulating or  
XX inhibiting the expression of genes and/or translation of transcripts, and  
XX as therapeutic agents. The proteins and peptides are useful as  
XX therapeutic, prognostic and diagnostic reagents for cancer. The present  
XX sequence is a human leukocyte antigen (HLA) peptide, used in an example  
XX from the invention  
XX  
XX Sequence 9 AA;  
XX  
XX Query Match 57.1%; Score 24; DB 6; Length 9;  
XX Best Local Similarity 55.6%; Pred. No. 1.7e+06;  
XX Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
Db 1 NLSDMKREV 9

RESULT 5  
ID ABR20369 standard; peptide; 9 AA.  
XX ABR20369;  
AC ABR20369;  
XX

DT 19-MAY-2003 (first entry)  
XX  
XX Human cancer-related protein 185P2C9 HLA peptide #204.  
XX  
XX Human; cytostatic; vaccine; cancer; immune response; HLA;  
XX human leukocyte antigen.  
XX Homo sapiens.  
XX WO200283921-A2.  
XX  
XX 24-OCT-2002.  
XX  
XX 10-APR-2002; 2002WO-US011654.  
XX  
XX 10-APR-2001; 2001US-0282739P.  
XX 10-APR-2001; 2001US-0283112P.  
XX 25-APR-2001; 2001US-0286630P.  
XX  
XX (AGEN-) AGENSYS INC.  
XX  
XX Jakobovits A, Chailita-Bid PM, Paris M, Ge W, Hubert RS;  
XX Morrison K, Morrison RK, Raitano AB;  
XX WPI; 2003-075555/07.  
XX  
XX New composition comprising a substance that modulates the structure of  
XX proteins and polynucleotides, useful for therapeutic, prognostic and  
XX diagnostic reagents for eliciting cellular or humoral immune response in  
XX cancer patients.  
XX  
XX Claim 13; Page 322; 1021pp; English.  
XX  
XX The present invention relates to novel human cancer-related genes and  
XX proteins (AB278120-AB278168 and ABR01789-ABR01861). The genes and  
XX proteins are useful for eliciting a humoral or cellular immune response.  
XX The genes are useful as probes and primers for the amplification and/or  
XX detection of genes, mRNAs or their fragments, as reagents for the  
XX diagnosis and/or prognosis of cancer, as coding sequences capable of  
XX directing the expression of the protein, as tools for modulating or  
XX inhibiting the expression of genes and/or translation of transcripts, and  
XX as therapeutic agents. The proteins and peptides are useful as  
XX therapeutic, prognostic and diagnostic reagents for cancer. The present  
XX sequence is a human leukocyte antigen (HLA) peptide, used in an example  
XX from the invention  
XX  
XX Sequence 9 AA;  
XX  
XX Query Match 57.1%; Score 24; DB 6; Length 9;  
XX Best Local Similarity 55.6%; Pred. No. 1.7e+06;  
XX Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLTANKEV 9  
Db 1 NLSDMKREV 9

RESULT 6  
ID ABJ19931 standard; peptide; 9 AA.  
XX ABJ19931;  
AC ABJ19931;  
XX  
XX 10-APR-2003 (first entry)  
XX  
XX MHC binding peptide SEQ ID No 96.  
XX  
XX Antiinflammatory; antiallergic; antiarthritic; nootropic; neuroprotective;  
XX antiinflammatory; major histocompatibility complex; MHC;  
XX autoimmune disease; T cell; B cell; allergic disease; multiple sclerosis;  
XX rheumatoid arthritis; neurodegenerative disorder; Alzheimer's disease;  
XX inflammation; gene therapy; MHC binding peptide.  
XX

OS Synthetic.  
XX WO200294981-A2.  
XX 28-NOV-2002.  
XX 16-MAY-2002; 2002WO-11000383.  
XX 16-MAY-2001; 2001US-02809587.  
XX 29-MAY-2001; 2001US-00865348.  
XX (TECR ) TECHNION RES & DEV FOUND LTD.  
XX Barnea E, Beer I, Ziv T, Admon A, Daessau L, Buchsbaum S;  
XX WPI; 2003-210043/20.  
XX  
XX Identifying peptides that are capable of binding to major  
XX histocompatibility complex (MHC) molecules of a particular haplotype by  
XX analyzing peptides bound to the soluble and secreted form of the MHC  
XX molecules of the particular haplotype.  
XX  
XX Claim 37; Page 169; 238pp; English.  
XX  
XX The invention relates to a novel method for identifying peptides  
XX originating from a particular cell type, which are capable of binding to  
XX major histocompatibility complex (MHC) molecules of a particular  
XX haplotype. The method comprises analyzing peptides bound to the soluble  
XX and secreted form of the MHC molecules of the particular haplotype. The  
XX method is useful for identifying peptides for treating an autoimmune  
XX disease, such as T or B cell and/or allergic disease or condition,  
XX rheumatoid arthritis, or multiple sclerosis, neurodegenerative disorders,  
XX e.g. Alzheimer's disease, or diseases associated with inflammation. The  
XX sequences of the invention may be used in a gene therapy application.  
XX This sequence represents a peptide relating to the method for identifying  
XX MHC binding peptides of the invention  
XX  
XX Sequence 9 AA;  
XX  
XX Query Match 54.8%; Score 23; DB 6; Length 9;  
XX Best Local Similarity 44.4%; Pred. No. 1.7e+06;  
XX Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
XX  
XX 1 SLTANMKEV 9  
XX :||:|:  
XX 1 ALTGHEEV 9  
XX  
XX  
XX RESULT 7  
XX ABJ20118  
XX ID ABJ20118 standard; peptide; 9 AA.  
XX  
XX AC ABJ20118;  
XX  
XX DT 10-APR-2003 (first entry)  
XX  
XX DE MHC binding peptide SEQ ID No 283.  
XX  
XX Antiinflammatory; antiallergic; antiarthritic; neuroprotective;  
XX autoinflammatory; major histocompatibility complex; MHC;  
XX autoimmune disease; T cell; B cell; allergic disease; multiple sclerosis;  
XX rheumatoid arthritis; neurodegenerative disorder; Alzheimer's disease;  
XX inflammation; gene therapy; MHC binding peptide.  
XX  
XX Synthetic.  
XX  
XX WO200294981-A2.  
XX 28-NOV-2002.  
XX 16-MAY-2002; 2002WO-11000383.  
XX 29-MAY-2001; 2001US-02809587.  
XX 16-MAY-2001; 2001US-02809587.

PR 29-MAY-2001; 2001US-00865548.  
XX  
XX (TECR ) TECHNION RES & DEV FOUND LTD.  
XX  
XX Barnea E, Beer I, Ziv T, Admon A, Daessau L, Buchsbaum S;  
XX WPI; 2003-210043/20.  
XX  
XX Identifying peptides that are capable of binding to major  
XX histocompatibility complex (MHC) molecules of a particular haplotype by  
XX analyzing peptides bound to the soluble and secreted form of the MHC  
XX molecules of the particular haplotype.  
XX  
XX Claim 58; Page 216; 238pp; English.  
XX  
XX The invention relates to a novel method for identifying peptides  
XX originating from a particular cell type, which are capable of binding to  
XX major histocompatibility complex (MHC) molecules of a particular  
XX haplotype. The method comprises analyzing peptides bound to the soluble  
XX and secreted form of the MHC molecules of the particular haplotype. The  
XX method is useful for identifying peptides for treating an autoimmune  
XX disease, such as T or B cell and/or allergic disease or condition,  
XX rheumatoid arthritis, or multiple sclerosis, neurodegenerative disorders,  
XX e.g. Alzheimer's disease, or diseases associated with inflammation. The  
XX sequences of the invention may be used in a gene therapy application.  
XX This sequence represents a peptide relating to the method for identifying  
XX MHC binding peptides of the invention  
XX  
XX Sequence 9 AA;  
XX  
XX Query Match 54.8%; Score 23; DB 6; Length 9;  
XX Best Local Similarity 55.6%; Pred. No. 1.7e+06;  
XX Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
XX  
XX 1 SLTANMKEV 9  
XX :||:|:  
XX 1 SLTANMLEI 9  
XX  
XX  
XX RESULT 8  
XX ABR23049  
XX ID ABR23049 standard; peptide; 9 AA.  
XX  
XX AC ABR23049;  
XX  
XX DT 19-MAY-2003 (first entry)  
XX  
XX DE Human cancer-related protein 185P2C9 HLA peptide #2884.  
XX  
XX Human; cytostatic; vaccine; cancer; immune response; HLA;  
XX human leukocyte antigen.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200283921-A2.  
XX  
XX PD 24-OCT-2002.  
XX  
XX 10-APR-2002; 2002WO-US011654.  
XX  
XX 10-APR-2001; 2001US-0282739P.  
XX 10-APR-2001; 2001US-0283112P.  
XX 25-APR-2001; 2001US-0286630P.  
XX  
XX (AGEN-) AGENSYS INC.  
XX  
XX Jakobovits A, Challita-Bid PM, Faris M, Ge W, Hubert RS;  
XX Morrison K, Morrison RK, Raitano AB;  
XX WPI; 2003-075555/07.  
XX  
XX New composition comprising a substance that modulates the structure of  
XX proteins and polynucleotides, useful for therapeutic, prognostic and  
XX

PT diagnostic reagents for eliciting cellular or humoral immune response in  
 PT cancer patients.  
 XX  
 PS Claim 13, Page 353, 1021p; English.  
 XX  
 CC The present invention relates to novel human cancer-related genes and  
 CC proteins (AB278120-AB278168 and ABR01789-ABR01861). The genes and  
 CC proteins are useful for eliciting a humoral or cellular immune response.  
 CC The genes are useful as probes and primers for the amplification and/or  
 CC detection of genes, mRNAs or their fragments, as reagents for the  
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of  
 CC directing the expression of the protein, as tools for modulating or  
 CC inhibiting the expression of genes and/or translation of transcripts, and  
 CC as therapeutic agents. The proteins and peptides are useful as  
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present  
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example  
 CC from the invention  
 CC  
 SQ Sequence 9 AA;  
 Query Match 54.8%; Score 23; DB 6; Length 9;  
 Best Local Similarity 62.5%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 LTANKKEV 9  
 : : : : :  
 1 LSDMKKEV 8  
 DB  
 RESULT 9  
 ABR20246  
 ID ABR20246 standard; peptide; 9 AA.  
 XX  
 AC ABR20246;  
 XX  
 DT 19-MAY-2003 (first entry)  
 XX  
 DE Human cancer-related protein 185P2C9 HLA peptide #91.  
 XX  
 KW Human; cytosolic; vaccine; cancer; immune response; HLA;  
 KW human leukocyte antigen.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200283921-A2.  
 XX  
 PD 24-OCT-2002.  
 XX  
 PF 10-APR-2002; 2002WO-US011654.  
 XX  
 PR 10-APR-2001; 2001US-0282739P.  
 PR 10-APR-2001; 2001US-0283112P.  
 PR 25-APR-2001; 2001US-0286630P.  
 XX  
 PA (AGEN-) AGENSYS INC.  
 XX  
 PI Jakovcivite A, Challita-Eid PM, Faris M, Ge W, Hubert RS;  
 PI Morrison K, Morrison RK, Raitano AB;  
 XX  
 DR WPI; 2003-075555/07.  
 XX  
 PT New composition comprising a substance that modulates the structure of  
 PT proteins and polynucleotides; useful for therapeutic, prognostic and  
 PT diagnostic reagents for eliciting cellular or humoral immune response in  
 PT cancer patients.  
 XX  
 PS Claim 13, Page 321; 1021p; English.  
 XX  
 CC The present invention relates to novel human cancer-related genes and  
 CC proteins (AB278120-AB278168 and ABR01789-ABR01861). The genes and  
 CC proteins are useful for eliciting a humoral or cellular immune response.  
 CC The genes are useful as probes and primers for the amplification and/or  
 CC detection of genes, mRNAs or their fragments, as reagents for the

CC diagnosis and/or prognosis of cancer, as coding sequences capable of  
 CC directing the expression of the protein, as tools for modulating or  
 CC inhibiting the expression of genes and/or translation of transcripts, and  
 CC as therapeutic agents. The proteins and peptides are useful as  
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present  
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example  
 CC from the invention  
 CC  
 SQ Sequence 9 AA;  
 Query Match 54.8%; Score 23; DB 6; Length 9;  
 Best Local Similarity 62.5%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 LTANKKEV 9  
 : : : : :  
 1 LSDMKKEV 8  
 DB  
 RESULT 10  
 ABR21637  
 ID ABR21637 standard; peptide; 9 AA.  
 XX  
 AC ABR21637;  
 XX  
 DT 19-MAY-2003 (first entry)  
 XX  
 DE Human cancer-related protein 185P2C9 HLA peptide #1472.  
 XX  
 KW Human; cytosolic; vaccine; cancer; immune response; HLA;  
 KW human leukocyte antigen.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200283921-A2.  
 XX  
 PD 24-OCT-2002.  
 XX  
 PF 10-APR-2002; 2002WO-US011654.  
 XX  
 PR 10-APR-2001; 2001US-0282739P.  
 PR 10-APR-2001; 2001US-0283112P.  
 PR 25-APR-2001; 2001US-0286630P.  
 XX  
 PA (AGEN-) AGENSYS INC.  
 XX  
 PI Jakovcivite A, Challita-Eid PM, Faris M, Ge W, Hubert RS;  
 PI Morrison K, Morrison RK, Raitano AB;  
 XX  
 DR WPI; 2003-075555/07.  
 XX  
 PT New composition comprising a substance that modulates the structure of  
 PT proteins and polynucleotides; useful for therapeutic, prognostic and  
 PT diagnostic reagents for eliciting cellular or humoral immune response in  
 PT cancer patients.  
 XX  
 PS Claim 13, Page 337; 1021p; English.  
 XX  
 CC The present invention relates to novel human cancer-related genes and  
 CC proteins (AB278120-AB278168 and ABR01789-ABR01861). The genes and  
 CC proteins are useful for eliciting a humoral or cellular immune response.  
 CC The genes are useful as probes and primers for the amplification and/or  
 CC detection of genes, mRNAs or their fragments, as reagents for the  
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of  
 CC directing the expression of the protein, as tools for modulating or  
 CC inhibiting the expression of genes and/or translation of transcripts, and  
 CC as therapeutic agents. The proteins and peptides are useful as  
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present  
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example  
 CC from the invention  
 CC  
 SQ Sequence 9 AA;

Query Match 54.8%; Score 23; DB 6; Length 9;  
 Best Local Similarity 62.5%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LTANKKEV 9  
 ID 1 LSDMKKEV 8

RESULT 11  
 ADH68371  
 ID ADH68371 standard; peptide; 6 AA.

XX ADH68371;  
 AC  
 XX 25-MAR-2004 (first entry)  
 XX  
 DE RSV F-protein peptide fragment #2.  
 XX  
 XX vaccine; metapneumovirus family; respiratory syncytial virus;  
 KM fusion protein; F protein; attenuated; stabilisation; virucide.  
 XX  
 OS Respiratory syncytial virus.

XX WO2003097089-A2.  
 XX 27-NOV-2003.  
 XX  
 PF 16-MAY-2003; 2003WO-EP005187.  
 XX  
 PR 16-MAY-2002; 2002DE-01021836.

XX (LOHM) LOHMANN ANIMAL HEALTH GMBH & CO KG.

XX Naylor CJ;  
 DR WPI; 2004-012486/01.

PT Vaccine against metapneumovirus or respiratory syncytial virus, useful  
 PT for preventing infections in birds or humans, contains virus with  
 PT modified fusion protein sequence.

XX Disclosure; SEQ ID NO 78; 100pp; German.

XX This invention describes a novel vaccine against a member of the  
 CC metapneumovirus family, respiratory syncytial virus (RSV) or a virus  
 CC having significant genetic homology with them in the fusion (F) protein  
 CC region, containing a virus (or part) in which the 293-296 amino acid  
 CC region of F protein (or functional equivalents, e.g. 323-328 in RSV) has  
 CC been modified. The preferred vaccine contains a live, preferably  
 CC attenuated, virus, and is formulated with auxiliary, carrier and/or  
 CC adjuvant, optionally also with interleukins 6 (particularly for avian  
 CC metapneumovirus) or 12 or 18 (particularly for human metapneumovirus or  
 CC RSV). The modification in the F protein is a stabilisation, especially  
 CC substitution of one codon by another that requires several mutations for  
 CC reversion to the wild type, specifically by a codon with low probability  
 CC of reversion to a glu-encoding codon. A virulent virus is modified in the  
 CC nucleic acid sequence that encodes the specified part of the F protein.  
 CC Particularly a full-length DNA copy of the genome is assembled by  
 CC ligating PCR products that include the required modification, then virus  
 CC is rescued using e.g. chickenpox IT polymerase or cellular ribosomal pol  
 CC RNA polymerase. The vaccines, also the modified F proteins, are used to  
 CC prevent disease caused by human or avian metapneumovirus or RSV.  
 CC Modification of the specified region of the F protein, which is important  
 CC for virulence, results in a stable attenuation, resistant to reversion to  
 CC the wild type.

XX Sequence 6 AA:

Query Match 52.4%; Score 22; DB 8; Length 6;  
 Best Local Similarity 66.7%; Pred. No. 1.7e+06;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 TANKE 8  
 ID 1 TTANKKE 6

RESULT 12  
 AAB46277  
 ID AAB46277 standard; peptide; 9 AA.

XX AAB46277;  
 AC  
 XX 04-APR-2001 (first entry)  
 XX

DE HPV type 16 L1 protein cytotoxic T-cell epitope SEQ ID NO 5.

XX Cytotoxic T cell epitope; L1 protein; antigenic; vaccine; antiviral;  
 KM antitumor; immune response; tumor.

XX Human papillomavirus.

XX DE19925199-A1.

XX 07-DEC-2000.

XX 01-JUN-1999; 99DE-01025199.

XX 01-JUN-1999; 99DE-01025199.

XX (MED1-) MEDIGENE AG.

XX Nieland J, Jochmus I;

XX WPI, 2001-072111/09.

PT New T cell epitopes from the papilloma virus L1 protein, useful for  
 PT detecting or inducing an immune response, e.g. in vaccines.

PS Claim 1; Page 13; 24pp; German.

XX This invention describes novel T cell epitopes (A) of the human papilloma  
 CC virus L1 protein. The invention also describes (a) compounds (I)  
 CC containing an (A) provided it is not a natural L1 protein or an  
 CC exclusively N- or C-terminal deletion variant of natural L1; (b) nucleic  
 CC acid (II) that encodes (A) or (I); (c) vector, particularly an expression  
 CC vector, that contains (II); (d) cells that contain, and preferably  
 CC present, (A); (e) complex of (A) or (I) with at least one additional  
 CC molecule (III); in vitro detection of T cell activation by a compound  
 CC containing at least one (A); (f) preparing cells of (d); and (g) test  
 CC system for in vitro detection of T cells. The products of the invention  
 CC have antiviral and antitumor activity. (A), also compounds or complexes  
 CC containing it or nucleic acid or vectors encoding it, is used for  
 CC detection of an immune response, particularly detecting L1-specific  
 CC cytotoxic T cells or to determine the L1 protein-specific antigenicity of  
 CC compounds and complexes that contain (A), and for inducing a protective  
 CC immune response against papilloma virus and related tumors, particularly  
 CC as vaccines

XX Sequence 9 AA:

Query Match 52.4%; Score 22; DB 4; Length 9;  
 Best Local Similarity 55.6%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LTANKKEV 9  
 ID 1 TTANKKEV 9

XX Sequence 9 AA:

RESULT 13  
 ADH40759  
 ID ADH40759 standard; peptide; 9 AA.  
 XX  
 AC ADH40759;

XX 11-MAR-2004 (first entry)  
XX Human CD42b HLA binding peptide.  
XX human; cytostatic; vaccine; SNP profile; cancer; leukaemia; HLA.  
XX Homo sapiens.  
XX WO2003106692-A2.  
XX 24-DEC-2003.  
XX 13-JUN-2003; 2003WO-EP006251.  
XX 13-JUN-2002; 2002EP-00013423.  
XX (MERCK) MERCK PATENT GMBH.  
XX Stritmatter W, Moll H;  
XX WPI; 2004-082200/08.  
XX Providing allelic variant epitope of protein based on single nucleotide  
PT polymorphism by defining target protein, screening database of protein,  
PT identifying, selecting allelic variant protein, creating variant  
PT epitopes.  
XX Disclosure; Page 96; 11pp; English.  
XX The invention relates to a novel method for providing epitopes of allelic  
CC variants of antigenic proteins from specific species based on single  
CC nucleotide polymorphism (SNP), by defining target protein/peptide or its  
CC subset, screening database of DNA encoding target protein, identifying,  
CC selecting allelic peptide/protein variants, expression product or its  
CC fragment encoded by DNA sequence having SNP, creating variant epitopes,  
CC selecting epitopes binding to MHC protein. A protein of the invention has  
CC cytostatic activity, and may have a use in a vaccine. The method is  
CC useful for generating a SNP profile of one or more individuals from a  
CC given species by applying the method for several protein from the  
CC individuals, where the SNP profile was related to disease, preferably  
CC cancer. This is useful for diagnosing a disease in an individual by  
CC generating the SNP-related polymorphic profile. A method of the invention  
CC is useful for transplanting haematopoietic stem cells from a donor to a  
CC recipient and treating cancer, preferably leukaemia, and for determining  
CC the progression, regression or onset of a treated disease. The present  
CC sequence is used in the exemplification of the invention.  
XX Sequence 9 AA;  
SQ  
Query Match 52.4%; Score 22; DB 8; Length 9;  
Best Local Similarity 44.4%; Pred. No. 1.7e+06;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
CY 1 SLTANKEV 9  
DB 1 SMISNVASV 9  
RESULT 14  
ADH40364  
ID ADH40364 standard; peptide; 9 AA.  
XX ADH40364;  
XX 11-MAR-2004 (first entry)  
XX Human CD42b HLA class I binding peptide.  
XX human; cytostatic; vaccine; SNP profile; cancer; leukaemia; HLA.  
XX Homo sapiens.  
XX

PN WO2003106692-A2.  
XX 24-DEC-2003.  
XX 13-JUN-2003; 2003WO-EP006251.  
XX 13-JUN-2002; 2002EP-00013423.  
XX (MERCK) MERCK PATENT GMBH.  
XX Stritmatter W, Moll H;  
XX WPI; 2004-082200/08.  
XX Providing allelic variant epitope of protein based on single nucleotide  
PT polymorphism by defining target protein, screening database of protein,  
PT identifying, selecting allelic variant protein, creating variant  
PT epitopes.  
XX Disclosure; Page 88; 11pp; English.  
XX The invention relates to a novel method for providing epitopes of allelic  
CC variants of antigenic proteins from specific species based on single  
CC nucleotide polymorphism (SNP), by defining target protein/peptide or its  
CC subset, screening database of DNA encoding target protein, identifying,  
CC selecting allelic peptide/protein variants, expression product or its  
CC fragment encoded by DNA sequence having SNP, creating variant epitopes,  
CC selecting epitopes binding to MHC protein. A protein of the invention has  
CC cytostatic activity, and may have a use in a vaccine. The method is  
CC useful for generating a SNP profile of one or more individuals from a  
CC given species by applying the method for several protein from the  
CC individuals, where the SNP profile was related to disease, preferably  
CC cancer. This is useful for diagnosing a disease in an individual by  
CC generating the SNP-related polymorphic profile. A method of the invention  
CC is useful for transplanting haematopoietic stem cells from a donor to a  
CC recipient and treating cancer, preferably leukaemia, and for determining  
CC the progression, regression or onset of a treated disease. The present  
CC sequence is used in the exemplification of the invention.  
XX Sequence 9 AA;  
SQ  
Query Match 52.4%; Score 22; DB 8; Length 9;  
Best Local Similarity 44.4%; Pred. No. 1.7e+06;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
CY 1 SLTANKEV 9  
DB 1 SMISNVASV 9  
RESULT 15  
AAG80232  
ID AAG80232 standard; peptide; 8 AA.  
XX AAG80232;  
XX 23-JAN-2002 (first entry)  
XX HOP protein TPRI domain associated oligopeptide #4.  
XX Hop protein; chaperone protein; TPR motif; tetratricco peptide repeat;  
XX three-dimensional structure; 3D; cytostatic; virucide; immunosuppressive;  
XX antiinflammatory; tumour treatment; autoimmune disease; TPR;  
XX protein coordinate data.  
XX Undentified.  
XX DE10018335-A1.  
XX 04-OCT-2001.  
XX 13-APR-2000; 2000DE-01018335.  
XX

PR 29-MAR-2000; 2000DE-01015748.

XX (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

XX Moarefi I, Scheufler C, Hartl U, Brinker A;

XX WFI; 2002-000454/01.

XX New spatial model of polypeptide containing a TPR motif, useful for  
PT identifying inhibitors of interaction between chaperone and Hop proteins,  
PT potential therapeutics.

XX Disclosure: Fig 6; 145pp; German.

CC This invention describes a novel spatial model of a polypeptide (I)  
CC having at least one amino acid (aa) sequence (II) of a TPR (tetratricic  
CC peptide repeat) structural motif of a Hop protein, or a derivative of  
CC (II). The invention also describes a method for (1) producing a crystal  
CC containing, per asymmetric unit of its unit cell, at least one (I) and  
CC optionally another compound; (2) preparing the crystals of (1); (3)  
CC preparing compounds (II) that bind, as ligands, to a structural region  
CC of a Hop protein (or structurally related protein) by non-covalent  
CC interaction with the main and/or side chains of aa that are components of  
CC the TPR domain; (4) preparing peptides (Ia), having the spatial model of  
CC (I), that can bind to a chaperone protein (CP) but are non-functional as  
CC regards adapter function; (5) identifying compounds (A) that inhibit  
CC interaction between CP and Hop proteins; (6) (A) identified by the method  
CC of (5); (7) representing the three-dimensional (3D) structure of a  
CC polypeptide of unknown structure but containing at least one TPR motif  
CC (or its derivative or complex), using the model of (1); (8) identifying  
CC compounds (A') that inhibit interaction between CP and polypeptides  
CC containing at least one TPR motif, or its derivatives; (9) pharmaceutical  
CC compositions containing (A) or (A'); (10) preparing DNA sequences (IV)  
CC that encode a sequence that contains at least part of proteins rct1 to 4,  
CC IRSF, SGT and KIAA0719; (11) preparing amino acid sequences (V) encoded  
CC by (IV); and (12) preparing pharmaceutical composition containing (V),  
CC their fragments or derivatives. The products of the invention have  
CC cytosstatic, virocidic, immunosuppressive and antiinflammatory activity.  
CC The models are used to identify compounds (A) that inhibit interaction  
CC between Hop and chaperone proteins, potentially useful for treating  
CC tumours and virus infections, also as immunosuppressives, for treating  
CC autoimmune diseases or other inflammatory states or guest versus host  
CC disease. This sequence represents a Hop protein TPR domain associated  
CC oligopeptide used to illustrate the method of the invention

XX SQ Sequence 8 AA;

Query Match 50.0%; Score 21; DB 5; Length 8;

Best Local Similarity 57.1%; Pred. No. 1.7e+06;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 TANKEV 9  
| | | | |  
Db 1 TSMREEV 7

Search completed: November 30, 2004, 08:33:08  
Job time : 117 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 08:34:19 ; Search time 101 Seconds  
(without alignments)  
31.606 Million cell updates/sec

Title: US-10-008-377A-7  
Perfect score: 42  
Sequence: 1 SLTANKREV 9.

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1575965 seqs, 354694765 residues

Total number of hits satisfying chosen parameters: 135309

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/ECT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	100.0	9	14 US-10-008-377-7	Sequence 7, Appl
2	24	57.1	9	9 US-09-835-948-84	Sequence 84, Appl
3	24	57.1	9	9 US-09-835-948-141	Sequence 141, Appl
4	24	57.1	9	14 US-10-172-597-84	Sequence 84, Appl
5	24	57.1	9	14 US-10-172-597-141	Sequence 141, Appl
6	23	54.8	9	10 US-09-865-548A-96	Sequence 96, Appl
7	20	47.6	8	14 US-10-008-524A-78	Sequence 78, Appl
8	20	47.6	8	14 US-10-008-524A-79	Sequence 79, Appl
9	20	47.6	8	14 US-10-008-524A-80	Sequence 80, Appl
10	20	47.6	8	14 US-10-350-719-78	Sequence 78, Appl
11	20	47.6	8	14 US-10-350-719-79	Sequence 79, Appl
12	20	47.6	8	14 US-10-350-719-80	Sequence 80, Appl
13	20	47.6	9	10 US-09-932-165-74	Sequence 74, Appl

14	20	47.6	9	10 US-09-932-165-160	Sequence 160, App
15	20	47.6	9	10 US-09-932-165-664	Sequence 664, App
16	20	47.6	9	14 US-10-334-726-109	Sequence 109, App
17	19	45.2	6	9 US-09-842-628-18	Sequence 18, Appl
18	19	45.2	6	9 US-09-842-628-22	Sequence 22, Appl
19	19	45.2	7	9 US-09-989-789-811	Sequence 811, App
20	19	45.2	7	10 US-09-990-186-811	Sequence 811, App
21	19	45.2	7	10 US-09-989-994-811	Sequence 811, App
22	19	45.2	7	17 US-10-478-991-67	Sequence 67, Appl
23	19	45.2	8	9 US-09-780-718-4	Sequence 4, Appl
24	19	45.2	8	9 US-09-949-972-8	Sequence 8, Appl
25	19	45.2	8	14 US-10-162-442A-8	Sequence 8, Appl
26	19	45.2	8	14 US-10-376-121A-154	Sequence 154, App
27	19	45.2	9	9 US-09-834-765-135	Sequence 135, App
28	19	45.2	9	9 US-09-834-765-249	Sequence 249, App
29	19	45.2	9	9 US-09-834-765-345	Sequence 345, App
30	19	45.2	9	9 US-09-834-765-441	Sequence 441, App
31	19	45.2	9	9 US-09-834-765-539	Sequence 539, App
32	19	45.2	9	14 US-10-084-813-902	Sequence 902, App
33	19	45.2	9	14 US-10-084-813-903	Sequence 903, App
34	19	45.2	9	14 US-10-084-813-904	Sequence 904, App
35	19	45.2	9	14 US-10-084-813-905	Sequence 905, App
36	19	45.2	9	14 US-10-084-813-906	Sequence 906, App
37	19	45.2	9	14 US-10-084-813-907	Sequence 907, App
38	19	45.2	9	14 US-10-119-536A-151	Sequence 151, App
39	19	45.2	9	14 US-10-119-536A-151	Sequence 151, App
40	19	45.2	9	14 US-10-245-871-703	Sequence 703, App
41	19	45.2	9	15 US-10-253-286-703	Sequence 703, App
42	18	42.9	8	14 US-10-193-754-75	Sequence 75, Appl
43	18	42.9	8	15 US-10-149-138-10	Sequence 10, Appl
44	18	42.9	8	15 US-10-149-138-691	Sequence 691, App
45	18	42.9	15	US-10-149-138-1979	Sequence 1979, App

## ALIGNMENTS

RESULT 1  
US-10-008-377-7  
Sequence 7, Application US/10008377  
Publication No. US20030157101A1  
GENERAL INFORMATION:  
APPLICANT: Gambacorti-Passerini, Carlo  
APPLICANT: Passoni, Lorena  
TITLE OF INVENTION: Immunogenic ALK Peptides  
FILE REFERENCE: 045922/241203  
CURRENT APPLICATION NUMBER: US/10/008,377  
CURRENT FILING DATE: 2001-11-15  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 7  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-008-377-7

Query Match  
Best Local Similarity 100.0%; Score 42; DB 14; Length 9;  
Pred. No. 1.4e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLTANKREV 9  
DB 1 SLTANKREV 9

RESULT 2  
US-09-835-948-84  
Sequence 84, Application US/09835948  
Patent No. US20020160365A1  
GENERAL INFORMATION:  
APPLICANT: O'Brien, Timothy J.  
TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of  
TITLE OF INVENTION: Ovarian Cancer

```
FILE REFERENCE: D6223CIP-B/D
CURRENT APPLICATION NUMBER: US/09/835,948
CURRENT FILING DATE: 2001-04-16
PRIOR APPLICATION NUMBER: US 09/492,543
PRIOR FILING DATE: 01-27-2000
NUMBER OF SEQ ID NOS: 189
SEQ ID NO 84
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Residues 54-62 of the PUMP-1 protein
US-09-835-948-84

Query Match
Best Local Similarity 57.1%; Score 24; DB 9; Length 9;
Pred. No. 1.4e+06;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTANKEV 9
DB 1 SLEAKLEKEM 9

RESULT 3
US-09-835-948-141
Sequence 141, Application US/09835948
Patent No. US20020160365A1
GENERAL INFORMATION:
APPLICANT: O'Brien, Timothy J.
TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
FILE REFERENCE: D6223CIP-B/D
CURRENT APPLICATION NUMBER: US/09/835,948
CURRENT FILING DATE: 2001-04-16
PRIOR APPLICATION NUMBER: US 09/492,543
PRIOR FILING DATE: 01-27-2000
NUMBER OF SEQ ID NOS: 189
SEQ ID NO 141
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Residues 54-62 of the PUMP-1 protein
US-09-835-948-141

Query Match
Best Local Similarity 57.1%; Score 24; DB 9; Length 9;
Pred. No. 1.4e+06;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTANKEV 9
DB 1 SLEAKLEKEM 9

RESULT 4
US-10-172-597-84
Sequence 84, Application US/10172597
Publication No. US20030180736A1
GENERAL INFORMATION:
APPLICANT: O'Brien, Timothy J.
TITLE OF INVENTION: Methods for the Early Diagnosis of Ovarian Cancer
FILE REFERENCE: D6223CIP-B/D/CIP
CURRENT APPLICATION NUMBER: US/10/172,597
CURRENT FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: US 09/835,948
PRIOR FILING DATE: 2001-04-16
NUMBER OF SEQ ID NOS: 189
SEQ ID NO 84
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Residues 54-62 of the PUMP-1 protein

US-10-172-597-84
Query Match
Best Local Similarity 57.1%; Score 24; DB 14; Length 9;
Pred. No. 1.4e+06;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTANKEV 9
DB 1 SLEAKLEKEM 9

RESULT 5
US-10-172-597-141
Sequence 141, Application US/10172597
Publication No. US20030180736A1
GENERAL INFORMATION:
APPLICANT: O'Brien, Timothy J.
TITLE OF INVENTION: Methods for the Early Diagnosis of Ovarian Cancer
FILE REFERENCE: D6223CIP-B/D/CIP
CURRENT APPLICATION NUMBER: US/10/172,597
CURRENT FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: US 09/835,948
PRIOR FILING DATE: 2001-04-16
NUMBER OF SEQ ID NOS: 189
SEQ ID NO 141
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Residues 54-62 of the PUMP-1 protein
US-10-172-597-141

Query Match
Best Local Similarity 57.1%; Score 24; DB 14; Length 9;
Pred. No. 1.4e+06;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLTANKEV 9
DB 1 SLEAKLEKEM 9

RESULT 6
US-09-865-548A-96
Sequence 96, Application US/0986548A
Publication No. US20030096298A1
GENERAL INFORMATION:
APPLICANT: Barnea, Elion
APPLICANT: Beer, Ilan
APPLICANT: Ziv, Tamir
APPLICANT: Admon, Aris
TITLE OF INVENTION: METHOD OF IDENTIFYING PEPTIDES CAPABLE OF BINDING TO MHC MOLECULE
FILE REFERENCE: 01/22080
CURRENT APPLICATION NUMBER: US/09/865,548A
CURRENT FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: US 60/290,958
PRIOR FILING DATE: 2001-05-16
NUMBER OF SEQ ID NOS: 204
SOFTWARE: PatentIn version 3.1
SEQ ID NO 96
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic peptide
US-09-865-548A-96

Query Match
Best Local Similarity 54.8%; Score 23; DB 10; Length 9;
Pred. No. 1.4e+06;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLTANKEV 9
DB 1 SLEAKLEKEM 9
```

Db 1 ALTGHEEV 9

## RESULT 7

US-10-008-524A-78  
; Sequence 78, Application US/10008524A  
; Publication No. US20030175682A1  
; GENERAL INFORMATION:  
; APPLICANT: Doobar, John  
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA  
; FILE REFERENCE: 18396/1074  
; CURRENT FILING DATE: 2002-08-13  
; PRIOR APPLICATION NUMBER: 09/314,268  
; PRIOR FILING DATE: 1999-05-18  
; NUMBER OF SEQ ID NOS: 179  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 78  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-008-524A-79

Query Match 47.6%; Score 20; DB 14; Length 8;

Best Local Similarity 57.1%; Pred. No. 1.4e+06; Indels 0; Gaps 0;

Matches 4; Conservative 2; Mismatches 1; Indels 0;

Qy 2 LTANKE 8

Db 2 LTAHKT 8

## RESULT 8

US-10-008-524A-79  
; Sequence 79, Application US/10008524A  
; Publication No. US20030175682A1  
; GENERAL INFORMATION:  
; APPLICANT: Doobar, John  
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA  
; FILE REFERENCE: 18396/1074  
; CURRENT APPLICATION NUMBER: US/10/008,524A  
; CURRENT FILING DATE: 2002-08-13  
; PRIOR APPLICATION NUMBER: 09/314,268  
; PRIOR FILING DATE: 1999-05-18  
; NUMBER OF SEQ ID NOS: 179  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 79  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-008-524A-79

Query Match 47.6%; Score 20; DB 14; Length 8;

Best Local Similarity 57.1%; Pred. No. 1.4e+06; Indels 0; Gaps 0;

Matches 4; Conservative 2; Mismatches 1; Indels 0;

Qy 2 LTANKE 8

Db 2 LTAHKT 8

## RESULT 9

US-10-008-524A-80  
; Sequence 80, Application US/10008524A  
; Publication No. US20030175682A1

; GENERAL INFORMATION:

; APPLICANT: Doobar, John  
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA  
; FILE REFERENCE: 18396/1074  
; CURRENT APPLICATION NUMBER: US/10/008,524A  
; CURRENT FILING DATE: 2002-08-13  
; PRIOR APPLICATION NUMBER: 09/314,268  
; PRIOR FILING DATE: 1999-05-18  
; NUMBER OF SEQ ID NOS: 179  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 80  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-008-524A-80

Query Match 47.6%; Score 20; DB 14; Length 8;

Best Local Similarity 57.1%; Pred. No. 1.4e+06; Indels 0; Gaps 0;

Matches 4; Conservative 2; Mismatches 1; Indels 0;

Qy 2 LTANKE 8

Db 1 LTAHKT 7

## RESULT 10

US-10-350-719-78  
; Sequence 78, Application US/10350719  
; Publication No. US20030219726A1  
; GENERAL INFORMATION:  
; APPLICANT: Doobar, John  
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA  
; FILE REFERENCE: 18396/2162  
; CURRENT APPLICATION NUMBER: US/10/350,719  
; CURRENT FILING DATE: 2003-01-24  
; PRIOR APPLICATION NUMBER: PCT/GB01/01176  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: GB0018140.4  
; PRIOR FILING DATE: 2000-07-24  
; NUMBER OF SEQ ID NOS: 180  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 78  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-350-719-78

Query Match 47.6%; Score 20; DB 14; Length 8;

Best Local Similarity 57.1%; Pred. No. 1.4e+06; Indels 0; Gaps 0;

Matches 4; Conservative 2; Mismatches 1; Indels 0;

Qy 2 LTANKE 8

Db 2 LTAHKT 8

## RESULT 11

US-10-350-719-79  
; Sequence 79, Application US/10350719  
; Publication No. US20030219726A1  
; GENERAL INFORMATION:  
; APPLICANT: Doobar, John  
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA  
; FILE REFERENCE: 18396/2162

```

; CURRENT APPLICATION NUMBER: US/10/350,719
; CURRENT FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: PCT/GB01/01176
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: GB0018140.4
; PRIOR FILING DATE: 2000-07-24
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 79
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-350-719-79

```

```

Query Match      47.6%; Score 20; DB 14; Length 8;
Best Local Similarity 57.1%; Pred. No. 1.4e+06;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2 LTANKKE 8
      |||:|:
Db      2 LTAHTKD 8

```

```

RESULT 12
US-10-350-719-80
; Sequence 80, Application US/10350719
; Publication No. US20030219726A1
; GENERAL INFORMATION:
; APPLICANT: Docrat, John
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
; FILE REFERENCE: 18396/2162
; CURRENT APPLICATION NUMBER: US/10/350,719
; CURRENT FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: PCT/GB01/01176
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: GB0018140.4
; PRIOR FILING DATE: 2000-07-24
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 80
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-350-719-80

```

```

Query Match      47.6%; Score 20; DB 14; Length 8;
Best Local Similarity 57.1%; Pred. No. 1.4e+06;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2 LTANKKE 8
      |||:|:
Db      1 LTAHTKD 7

```

```

RESULT 13
US-09-932-165-74
; Sequence 74, Application US/09932165
; Publication No. US20030134784A1
; GENERAL INFORMATION:
; APPLICANT: RAITANO, ARTHUR
; APPLICANT: CHALLITA-EID, PIA M.
; APPLICANT: FARIS, MARY
; APPLICANT: SAFERAN, DOUGLAS
; APPLICANT: AEFAR, DANIEL
; APPLICANT: LEVIN, ELANA
; APPLICANT: HUBERT, RENE

```

```

; APPLICANT: GE, WANGMAO
; APPLICANT: JAKOBOVITS, AVA
; TITLE OF INVENTION: NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED
; TITLE OF INVENTION: 83P2H3 AND CATP2E11 USEFUL IN TREATMENT AND
; TITLE OF INVENTION: DETECTION OF CANCER
; FILE REFERENCE: 51158-20014.00
; CURRENT APPLICATION NUMBER: US/09/932,165
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/226,329
; NUMBER OF SEQ ID NOS: 1508
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 74
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide motif
US-09-932-165-74

```

```

Query Match      47.6%; Score 20; DB 10; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      4 ANMK 7
      |||:|:
Db      6 ANMK 9

```

```

RESULT 14
US-09-932-165-460
; Sequence 460, Application US/09932165
; Publication No. US20030134784A1
; GENERAL INFORMATION:
; APPLICANT: RAITANO, ARTHUR
; APPLICANT: CHALLITA-EID, PIA M.
; APPLICANT: FARIS, MARY
; APPLICANT: SAFERAN, DOUGLAS
; APPLICANT: AEFAR, DANIEL
; APPLICANT: LEVIN, ELANA
; APPLICANT: HUBERT, RENE
; APPLICANT: GE, WANGMAO
; APPLICANT: JAKOBOVITS, AVA
; TITLE OF INVENTION: NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED
; TITLE OF INVENTION: 83P2H3 AND CATP2E11 USEFUL IN TREATMENT AND
; TITLE OF INVENTION: DETECTION OF CANCER
; FILE REFERENCE: 51158-20014.00
; CURRENT APPLICATION NUMBER: US/09/932,165
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/226,329
; NUMBER OF SEQ ID NOS: 1508
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 460
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide motif
US-09-932-165-460

```

```

Query Match      47.6%; Score 20; DB 10; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      4 ANMK 7
      |||:|:
Db      6 ANMK 9

```

```

RESULT 15
US-09-932-165-664
; Sequence 664, Application US/09932165

```

Publication No. US2003013478A1  
GENERAL INFORMATION:  
APPLICANT: RAITANO, ARTHUR  
APPLICANT: CHALLITA-ETD, PIA M.  
APPLICANT: FARIS, MARY  
APPLICANT: SAFERAN, DOUGLAS  
APPLICANT: AFEAR, DANIEL  
APPLICANT: LEVIN, ELANA  
APPLICANT: HUBERT, RENE  
APPLICANT: GE, WANGYAO  
APPLICANT: JAKOBOWITS, AYA  
TITLE OF INVENTION: NUCLEOTIC ACIDS AND CORRESPONDING PROTEINS ENTITLED  
TITLE OF INVENTION: 83P2H3 AND CATF2E11 USEFUL IN TREATMENT AND  
TITLE OF INVENTION: DETECTION OF CANCER  
FILE REFERENCE: 5.1158-20014.00  
CURRENT APPLICATION NUMBER: US/09/932,165  
CURRENT FILING DATE: 2001-08-17  
PRIOR APPLICATION NUMBER: 60/226,329  
PRIOR FILING DATE: 2000-08-17  
NUMBER OF SEQ ID NOS: 1508  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 664  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Peptide motif  
US-09-932-165-664

Query Match 47.6%; Score 20; DB 10; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ANMK 7  
DB 6 ANMK 9

Search completed: November 30, 2004, 08:41:16  
Job time : 101 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 08:28:43 ; Search time 20 Seconds  
(without alignments)  
29.843 Million cell updates/sec

Title: US-10-008-377A-7  
Perfect score: 42  
Sequence: 1 SLTANKEV 9

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 95011

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Issued Patents, AA:  
1: /cgn2\_6/prodata/1/aa/5A.COMB.pep.\*  
2: /cgn2\_6/prodata/1/aa/5B.COMB.pep.\*  
3: /cgn2\_6/prodata/1/aa/6A.COMB.pep.\*  
4: /cgn2\_6/prodata/1/aa/6B.COMB.pep.\*  
5: /cgn2\_6/prodata/1/aa/6C.COMB.pep.\*  
6: /cgn2\_6/prodata/1/aa/6D.COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	57.1	9	US-09-492-543-84	Sequence 84, Appl
2	57.1	9	US-09-492-543-141	Sequence 141, Appl
3	47.6	8	US-08-144-779C-31	Sequence 31, Appl
4	47.6	8	US-09-314-268-78	Sequence 78, Appl
5	47.6	8	US-09-314-268-79	Sequence 79, Appl
6	47.6	8	US-09-314-268-80	Sequence 80, Appl
7	47.6	8	US-09-368-449B-31	Sequence 31, Appl
8	47.6	9	US-08-144-779C-27	Sequence 27, Appl
9	47.6	9	US-09-368-449B-27	Sequence 27, Appl
10	45.2	6	US-08-050-482A-18	Sequence 18, Appl
11	45.2	6	US-08-050-482A-22	Sequence 22, Appl
12	45.2	8	US-08-718-905-8	Sequence 8, Appl
13	45.2	8	US-08-160-604-53	Sequence 53, Appl
14	45.2	8	US-09-550-497-8	Sequence 8, Appl
15	45.2	8	US-09-147-992-8	Sequence 8, Appl
16	45.2	8	US-09-549-839A-8	Sequence 8, Appl
17	45.2	8	US-08-475-955-154	Sequence 154, Appl
18	45.2	8	US-09-780-718-4	Sequence 4, Appl
19	45.2	9	US-09-357-293-4	Sequence 4, Appl
20	45.2	3	US-09-647-378B-48	Sequence 48, Appl
21	45.2	6	US-09-008-308-44	Sequence 44, Appl
22	42.9	6	US-09-256-237-44	Sequence 44, Appl
23	42.9	7	US-08-408-858A-55	Sequence 55, Appl
24	42.9	7	US-09-017-043A-2	Sequence 2, Appl
25	42.9	8	US-08-408-858A-61	Sequence 61, Appl
26	42.9	8	US-08-408-858A-62	Sequence 62, Appl
27	42.9	8	US-09-248-588-66	Sequence 66, Appl

28	18	42.9	8	4	US-09-206-947-79	Sequence 79, Appl
29	18	42.9	8	4	US-09-239-043D-583	Sequence 583, Appl
30	18	42.9	8	4	US-09-239-043D-599	Sequence 599, Appl
31	18	42.9	8	4	US-09-239-043D-1874	Sequence 1874, Appl
32	18	42.9	8	4	US-09-239-043D-1910	Sequence 1910, Appl
33	18	42.9	9	2	US-08-408-858A-68	Sequence 68, Appl
34	18	42.9	9	2	US-08-408-858A-69	Sequence 69, Appl
35	18	42.9	9	4	US-09-239-043D-353	Sequence 353, Appl
36	18	42.9	9	4	US-09-239-043D-600	Sequence 600, Appl
37	18	42.9	9	4	US-09-239-043D-868	Sequence 868, Appl
38	18	42.9	9	4	US-09-239-043D-1425	Sequence 1425, Appl
39	18	42.9	9	4	US-09-239-043D-1911	Sequence 1911, Appl
40	18	42.9	9	4	US-09-239-043D-1981	Sequence 1981, Appl
41	18	42.9	9	4	US-09-239-043D-2249	Sequence 2249, Appl
42	17	40.5	5	6	5217869-93	Patent No. 5217869
43	17	40.5	6	1	US-08-387-156-14	Sequence 14, Appl
44	17	40.5	6	1	US-08-742-256-10	Sequence 10, Appl
45	17	40.5	6	2	US-08-694-865-14	Sequence 14, Appl

## ALIGNMENTS

RESULT 1  
US-09-492-543-84  
Sequence 84, Application US/09492543A  
Patent No. 6316213  
GENERAL INFORMATION:  
APPLICANT: C'Brien, Timothy J.  
TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of  
FILE REFERENCE: D6223CIP-B  
CURRENT APPLICATION NUMBER: US/09/492,543A  
CURRENT FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: 09/039,211  
PRIOR FILING DATE: 03-14-1998  
NUMBER OF SEQ ID NOS: 189  
SOFTWARE: WORD 6.0.1 for Macintosh  
SEQ ID NO 84  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Residues 54-62 of the PUMP-1 protein  
US-09-492-543-84

Query Match 57.1%; Score 24; DB 3; Length 9;  
Best Local Similarity 55.6%; Pred. No. 3.8e+05;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 SLTANKEV 9  
DB 1 SLTANKEV 9  
1 SLTANKEV 9  
RESULT 2  
US-09-492-543-141  
Sequence 141, Application US/09492543A  
Patent No. 6316213  
GENERAL INFORMATION:  
APPLICANT: C'Brien, Timothy J.  
TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of  
FILE REFERENCE: D6223CIP-B  
CURRENT APPLICATION NUMBER: US/09/492,543A  
CURRENT FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: 09/039,211  
PRIOR FILING DATE: 03-14-1998  
NUMBER OF SEQ ID NOS: 189  
SOFTWARE: WORD 6.0.1 for Macintosh  
SEQ ID NO 141  
LENGTH: 9  
TYPE: PRT

ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Residues 54-62 of the PUMP-1 protein  
US-09-492-543-141

Query Match 57.1%; Score 24; DB 3; Length 9;  
Best Local Similarity 55.6%; Pred. No. 3.8e+05;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LITANKEV 9  
DB 1 LITANKEV 9

RESULT 3  
US-08-144-779C-31  
Sequence 31, Application US/08144779C  
Patent No. 6066822  
GENERAL INFORMATION:  
APPLICANT: Green, Lawrence R.  
APPLICANT: Sinackevich, Nicolay V.  
APPLICANT: Ivanov, Vadim T.  
APPLICANT: Mikhailova, Inessa I.  
APPLICANT: Vaskovsky, Boris V.  
APPLICANT: Mikhailov, Alexander N.  
APPLICANT: Khavinson, Vladimir K.  
APPLICANT: Morozov, Vyacheslav G.  
TITLE OF INVENTION: Immunomodulating Peptides and Methods  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08144,779C  
FILING DATE: 28-OCT-1993  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/783,517  
FILING DATE: 28-OCT-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/816,205  
FILING DATE: 02-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/967,633  
FILING DATE: 28-OCT-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Storella, John R.  
REGISTRATION NUMBER: 32,944  
REFERENCE/DOCKET NUMBER: 15542-S-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-144-779C-31

Query Match 47.6%; Score 20; DB 3; Length 8;  
Best Local Similarity 62.5%; Pred. No. 3.8e+05;

Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 LITANKEV 9  
DB 1 LITANKEV 8

RESULT 4  
US-09-314-268-78  
Sequence 78, Application US/09314268  
Patent No. 6346377  
GENERAL INFORMATION:  
APPLICANT: Doobatz, John  
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA  
FILE REFERENCE: 3789/80902  
CURRENT APPLICATION NUMBER: US/09/314,268  
CURRENT FILING DATE: 1999-03-19  
EARLIER APPLICATION NUMBER: 09/314,268  
EARLIER FILING DATE: 1999-05-18  
NUMBER OF SEQ ID NOS: 179  
SOFTWARE: Patent Ver. 2.1  
SEQ ID NO 78  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
OTHER INFORMATION: octapeptide antigen  
US-09-314-268-78

Query Match 47.6%; Score 20; DB 3; Length 8;  
Best Local Similarity 57.1%; Pred. No. 3.8e+05;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 LITANKEV 8  
DB 2 LITANKEV 8

RESULT 5  
US-09-314-268-79  
Sequence 79, Application US/09314268  
Patent No. 6346377  
GENERAL INFORMATION:  
APPLICANT: Doobatz, John  
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA  
FILE REFERENCE: 3789/80902  
CURRENT APPLICATION NUMBER: US/09/314,268  
CURRENT FILING DATE: 1999-03-19  
EARLIER APPLICATION NUMBER: 09/314,268  
EARLIER FILING DATE: 1999-05-18  
NUMBER OF SEQ ID NOS: 179  
SOFTWARE: Patent Ver. 2.1  
SEQ ID NO 79  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
OTHER INFORMATION: octapeptide antigen  
US-09-314-268-79

Query Match 47.6%; Score 20; DB 3; Length 8;  
Best Local Similarity 57.1%; Pred. No. 3.8e+05;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 LITANKEV 8  
DB 2 LITANKEV 8

RESULT 6  
US-09-314-268-80  
Sequence 80, Application US/09314268  
Patent No. 6346377  
GENERAL INFORMATION:  
APPLICANT: Doocbar, John  
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA  
FILE REFERENCE: 3789/80902  
CURRENT APPLICATION NUMBER: US/09/314,268  
CURRENT FILING DATE: 1999-03-19  
EARLIER APPLICATION NUMBER: 09/314,268  
EARLIER FILING DATE: 1999-05-18  
NUMBER OF SEQ. ID NOS: 179  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 80  
LENGTH: 8  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-314-268-80

Query Match 47.6%; Score 20; DB 3; Length 8;  
Best Local Similarity 57.1%; Pred. No. 3.8e+05;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

CY 2 LTANKE 8  
DB 1 LTAHMD 7

RESULT 7  
US-09-368-449B-31  
Sequence 31, Application US/09368449B  
Patent No. 6346514  
GENERAL INFORMATION:  
APPLICANT: Green, Lawrence R.  
Sinnackevich, Nicolay V.  
Ivanov, Vadim T.  
Mikhalyova, Inessa I.  
Vaskovsky, Boris V.  
Mikhaltsov, Alexander N.  
Khavinson, Vladimir K.  
Morozov, Vyacheslav G.  
TITLE OF INVENTION: Immunomodulating Peptides and Methods  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/368,449B  
FILING DATE: 04-Aug-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/144,779  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 07/816,205  
FILING DATE: 02-JAN-1992  
APPLICATION NUMBER: US 07/967,633  
FILING DATE: 28-OCT-1992  
ATTORNEY/AGENT INFORMATION:

NAME: Storella, John R.  
REGISTRATION NUMBER: 32,944  
REFERENCE/DOCKET NUMBER: 15542-5-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 31:  
US-09-368-449B-31

Query Match 47.6%; Score 20; DB 3; Length 8;  
Best Local Similarity 62.5%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

CY 2 LTANKEV 9  
DB 1 LTAZKAV 8

RESULT 8  
US-08-144-779C-27  
Sequence 27, Application US/08144779C  
Patent No. 606622  
GENERAL INFORMATION:  
APPLICANT: Green, Lawrence R.  
Sinnackevich, Nicolay V.  
Ivanov, Vadim T.  
Mikhalyova, Inessa I.  
Vaskovsky, Boris V.  
Mikhaltsov, Alexander N.  
Khavinson, Vladimir K.  
Morozov, Vyacheslav G.  
TITLE OF INVENTION: Immunomodulating Peptides and Methods  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/144,779C  
FILING DATE: 28-OCT-1993  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/783,517  
FILING DATE: 28-OCT-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/816,205  
FILING DATE: 02-JAN-1992  
APPLICATION NUMBER: US 07/967,633  
FILING DATE: 28-OCT-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Storella, John R.  
REGISTRATION NUMBER: 32,944  
REFERENCE/DOCKET NUMBER: 15542-5-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300

;; INFORMATION FOR SEQ ID NO: 27;  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 9 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS:  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-144-779C-27

Query Match 47.6%; Score 20; DB 3; Length 9;  
Best Local Similarity 62.5%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 LITANKEV 9  
||| |  
DB 2 LITAZKAV 9

RESULT 9  
US-09-368-449B-27  
; Sequence 27, Application US/09368449B  
; Patent No. 6346514  
; GENERAL INFORMATION:  
; APPLICANT: Green, Lawrence R.  
; Sinackevich, Nicolay V.  
; Ivanov, Vadim T.  
; Mikhailova, Inessa I.  
; Vaskovsky, Boris V.  
; Mikhailsov, Alexander N.  
; Khavinson, Vladimir K.  
; Morozov, Vyacheslav G.  
; TITLE OF INVENTION: Immunomodulating Peptides and Methods  
; of Use  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/368,449B  
; FILING DATE: 04-Aug-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/144,779  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US 07/816,205  
; FILING DATE: 02-JAN-1992  
; APPLICATION NUMBER: US 07/967,633  
; FILING DATE: 28-OCT-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Storella, John R.  
; REGISTRATION NUMBER: 32,944  
; REFERENCE/DOCKET NUMBER: 15542-5-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:  
US-09-368-449B-27

Query Match 47.6%; Score 20; DB 3; Length 9;  
Best Local Similarity 62.5%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 LITANKEV 9  
||| |  
DB 2 LITAZKAV 9

RESULT 10  
US-08-050-482A-18  
; Sequence 18, Application US/08050482A  
; Patent No. 6312939  
; GENERAL INFORMATION:  
; APPLICANT: ROBERTS, Joseph  
; MACALLISTER, Thomas W.  
; SETHURAMAN, Natarajan  
; FREEMAN, Abbie G.  
; TITLE OF INVENTION: GENETICALLY ENGINEERED GLUTAMINASE AND  
; ITS USE IN ANTIVIRAL AND ANTICANCER THERAPY  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY & LARDNER  
; STREET: 3000 K Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/050,482A  
; FILING DATE: 25-Apr-1995  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/US92/10421  
; FILING DATE: 04-DEC-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bent, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 023032/0106  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 672-5399  
; TELEFAX: (202) 672-5399  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:  
US-08-050-482A-18

Query Match 45.2%; Score 19; DB 3; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 MKEV 9  
||| |  
DB 1 MKEV 4

RESULT 11  
US-08-050-482A-22  
; Sequence 22, Application US/08050482A  
; Patent No. 6312939  
; GENERAL INFORMATION:  
; APPLICANT: ROBERTS, Joseph  
; MACALLISTER, Thomas W.

```

;
; SEHURAMAN, Natarajan
; FREEMAN, Abbie G.
; TITLE OF INVENTION: GENETICALLY ENGINEERED GLUTAMINASE AND
; ITS USE IN ANTIVIRAL AND ANTICANCER THERAPY
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/050,482A
; FILING DATE: 25-Apr-1995
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US92/10421
; FILING DATE: 04-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 25,768
; REFERENCE/DOCKET NUMBER: 023032/0106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-08-050-482A-22

Query Match 45.2%; Score 19; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 3.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 6 MKEY 9
DB 1 MKEY 4

RESULT 12
US-08-718-905-8
; Sequence 8, Application US/08718905
; Patent No. 6063756
; GENERAL INFORMATION:
; APPLICANT: Donovan, William P.
; APPLICANT: Donovan, Judith C.
; APPLICANT: Slaney, Annette C.
; TITLE OF INVENTION: BACILLUS THURINGIENSIS CRYET33 AND CRYET34
; TITLE OF INVENTION: COMPOSITIONS AND USES THEREFOR
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

```

;
; APPLICATION NUMBER: US/08/718,905
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: MOBT:003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-718-905-8

Query Match 45.2%; Score 19; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 3.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 6 MKEY 9
DB 2 MKEY 5

RESULT 13
US-08-160-604-53
; Sequence 53, Application US/08160604
; Patent No. 6232522
; GENERAL INFORMATION:
; APPLICANT: Harley, John
; APPLICANT: James, Judith A.
; APPLICANT: Scofield, R. H.
; TITLE OF INVENTION: PEPTIDE INDUCTION OF AUTOIMMUNITY AND CLINICAL SYMPTOMATOLOGY
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30308-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/160,604
; FILING DATE: 30-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/867,819
; FILING DATE: 13-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/648,205
; FILING DATE: 31-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/472,947
; FILING DATE: 31-JAN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRE114CIP(3)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)-815-6558
; TELEFAX: (404)-815-6555
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid

```

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
US-08-160-604-53

Query Match 45.2%; Score 19; DB 3; Length 8;  
Best Local Similarity 50.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LTKANKEV 9  
DB 1 LTKGKKEV 8

RESULT 14  
US-09-550-497-8

Sequence 8, Application US/09550497  
Patent No. 6248536

GENERAL INFORMATION:

APPLICANT: Donovan, William P.

Slaney, Judith C.

Donovan, Judith C.

TITLE OF INVENTION: BACILLUS THURINGIENSIS CRYET33 AND CRYET34

COMPOSITIONS AND USES THEREFOR

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: Texas

COUNTRY: United States of America

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/550,497

FILING DATE: 14-Apr-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/718,905

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Kitchell, Barbara S.

REGISTRATION NUMBER: 33,928

REFERENCE/DOCKET NUMBER: WOBET.003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000

TELEFAX: (512) 474-7577

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 8 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 8:

US-09-550-497-8

Query Match 45.2%; Score 19; DB 3; Length 8;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 MKEV 9  
DB 2 MKEV 5

RESULT 15

US-09-147-992-8

Sequence 8, Application US/09147992

Patent No. 6326351

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: BACILLUS THURINGIENSIS CRYET33 AND CRYET34

COMPOSITIONS AND USES THEREFOR

NUMBER OF SEQUENCES: 10

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/147,992

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/718,905

FILING DATE: 24-SEP-1996

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 8 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-09-147-992-8

Query Match 45.2%; Score 19; DB 3; Length 8;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 MKEV 9  
DB 2 MKEV 5

Search completed: November 30, 2004, 08:35:48  
Job time : 20 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: November 30, 2004, 07:41:32 / Search time 11.6667 Seconds  
(without alignments)  
82.472 Million cell updates/sec

Title: US-10-008-377a-6  
Percent score: 52  
Sequence: 1 ALPIEXGPLV 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 1102

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	38.5	7	2 A44428	platelet aggregati
2	18	34.6	9	2 S63491	dissimulatory sulf
3	18	34.6	9	2 S70332	endospore protein,
4	17	32.7	8	2 PT0311	Ig heavy chain CRD
5	17	32.7	8	2 E47393	neuropeptide calla
6	17	32.7	8	2 A28719	thymic humoral fac
7	17	32.7	9	2 G56978	collagen alpha 1(I
8	17	32.7	10	2 PC2171	triacylglycerol 11
9	16	30.8	5	2 UT0520	Ig kappa chain V-I
10	16	30.8	7	2 I48086	DNA topoisomerase
11	16	30.8	8	2 S15422	adipokinetic hormo
12	16	30.8	8	2 A58641	adipokinetic hormo
13	16	30.8	8	2 PT0368	Ig gamma chain C r
14	16	30.8	8	2 S21653	neuropeptide - flo
15	16	30.8	10	2 S65388	Cytochrome-c oxida
16	16	30.8	10	2 C39398	Fc mu (IgM) recept
17	16	30.8	10	2 PH1633	Ig H chain V-D-J r
18	15.5	29.8	10	2 C39111	Ig heavy chain C r
19	15	28.8	5	2 UN0860	peptidyl-dipeptida
20	15	28.8	6	2 A61049	halo-toxin - Pseud
21	15	28.8	7	2 S66442	glutathione S-tran
22	15	28.8	7	2 B48394	major fat-globule
23	15	28.8	8	2 D47393	neuropeptide calla
24	15	28.8	9	2 PT0285	Ig heavy chain CRD
25	15	28.8	9	2 S36850	Ig heavy chain V r
26	15	28.8	9	2 PT0080	60K Ca binding pro
27	15	28.8	9	2 PD0443	3-oxoacid CoA-tran
28	15	28.8	10	1 XAS8PC	angiotensin-conver
29	15	28.8	10	1 GMROL2	leucosulfakinin-II

30	15	28.8	10	2 B55482	platelet activatin
31	15	28.8	10	2 B60556	leucosulfakinin II
32	15	28.8	10	2 B59030	lyso-bradykinin -
33	14.5	27.9	9	2 G58502	kidney and bladder
34	14	26.9	5	2 B60274	major protein anti
35	14	26.9	5	2 PQ0689	photosystem I 10.4
36	14	26.9	7	2 I48105	dihydrofolate redu
37	14	26.9	7	2 EGMUCR	catch-relaxing pep
38	14	26.9	8	2 C39650	neural cell adhesi
39	14	26.9	9	2 S66398	quinoline-2-oxidor
40	14	26.9	9	2 B49712	ATP-binding protei
41	14	26.9	9	2 S26508	collagen alpha 2(V
42	14	26.9	9	2 B60246	ornitho-kinin - ch
43	14	26.9	10	1 RHLMGS	gonadoliberin - se
44	14	26.9	10	2 C26997	unspecific monooxy
45	14	26.9	10	2 P50451	24K protein 4302 -

#### ALIGNMENTS

RESULT 1  
A44428  
platelet aggregation-associated protein - Streptococcus sanguis (fragment)  
C:Species: Streptococcus sanguis  
C:Date: 11-Nov-1994 #sequence\_revision 11-Nov-1994 #text\_change 11-Nov-1994  
C:Accession: A44428  
R:Ericsson, P.R.; Herzberg, M.C.  
J. Biol. Chem. 268, 1646-1649, 1993  
A:Title: The Streptococcus sanguis platelet aggregation-associated protein. Identificat:  
A:Reference number: A44428; MUID:91131902; PMID:8420939  
A:Accession: A44428  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 17 <ERI>  
C:Keywords: cell wall

Query Match  
Best Local Similarity 38.5%; Score 20; DB 2; Length 7;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 2  
S63491  
dissimulatory sulfite reductase beta chain, membrane-bound - Desulfovibrio desulfuricans  
C:Species: Desulfovibrio desulfuricans  
C:Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 17-Mar-1999  
C:Accession: S63491  
R:Stender, J.; Arendsen, A.F.; Hagen, W.R.; Kroneck, P.M.H.  
Eur. J. Biochem. 233, 873-879, 1995  
A:Title: Molecular properties of the dissimulatory sulfite reductase from Desulfovibrio  
A:Reference number: S63491; MUID:96085152; PMID:8521853  
A:Accession: S63491  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-9 <STR>

Query Match  
Best Local Similarity 34.6%; Score 18; DB 2; Length 9;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 PIEVGP 8  
DB 4 PFGVNP 9

RESULT 3  
S70332  
endospore protein, 10K - rye (fragment)

C:Species: Secale cereale (rye)  
 C:Date: 19-Mar-1998 #sequence\_revision 17-Apr-1998 #text\_change 17-Apr-1998  
 C:Accession: S70332  
 R:Rocher, A.; Calero, M.; Soriano, F.; Mendez, E.  
 Biochim. Biophys. Acta 1295, 13-22, 1996  
 A:Title: Identification of major rye secalins as coeliac immunoreactive proteins.  
 A:Reference number: S70327; MUID:56283789; PMID:8679669  
 A:Accession: S70332  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-9 <RCC>

Query Match 34.6%; Score 18; DB 2; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 PLEV 6  
 ||||  
 DB 3 PLEV 6

RESULT 4  
 P70311  
 Ig heavy chain CDR3 region (clone 6-100) - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
 C:Accession: P70311  
 R:Yamada, M.; Masserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
 J. Exp. Med. 173, 395-407, 1991  
 A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and J  
 A:Reference number: P70222; MUID:91108337; PMID:1899102  
 A:Accession: P70311  
 A:Molecule type: DNA  
 A:Residues: 1-8 <YAM>  
 A:Experimental source: B lymphocyte  
 C:Keywords: heterodimer; immunoglobulin

Query Match 32.7%; Score 17; DB 2; Length 8;  
 Best Local Similarity 60.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 PLEV 7  
 ||||  
 DB 1 PSAYG 5

RESULT 5  
 E47393  
 neuropeptide callatostatin 5 - bluebottle fly (Calliphora vomitoria)  
 C:Species: Calliphora vomitoria  
 C:Date: 16-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
 C:Accession: E47393  
 R:Duve, H.; Johnson, A.H.; Scott, A.G.; Yu, C.G.; Yagci, K.J.; Tobe, S.S.; Thorpe, A.  
 Proc. Natl. Acad. Sci. U.S.A. 90, 2456-2460, 1993  
 A:Title: Callatostatin: neuropeptides from the blowfly Calliphora vomitoria with sequen  
 A:Reference number: A47393; MUID:93211980; PMID:8460157  
 A:Accession: E47393  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-8 <DUV>  
 A:Cross-references: UNIPROT:P41841  
 A:Experimental source: whole flies  
 A:Note: sequence extracted from NCBI backbone (NCBIP.128482)

Query Match 32.7%; Score 17; DB 2; Length 8;  
 Best Local Similarity 40.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 PLEV 7  
 ||||  
 DB 3 PYDFG 7

RESULT 6  
 A28719  
 thymic humoral factor gamma-2 - bovine (fragment)  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 09-Jul-2004  
 C:Accession: A28719  
 R:Birstein, Y.; Buchner, V.; Pecht, M.; Trainin, N.  
 Biochemistry 27, 4066-4071, 1988  
 A:Title: Thymic humoral factor gamma-2: purification and amino acid sequence of an immu  
 A:Reference number: A28719; MUID:88326920; PMID:3261994  
 A:Accession: A28719  
 A:Molecule type: protein  
 A:Residues: 1-8 <BUR>  
 A:Cross-references: UNIPROT:O18787

Query Match 32.7%; Score 17; DB 2; Length 8;  
 Best Local Similarity 60.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 IEXGP 8  
 ||||  
 DB 1 LEDGP 5

RESULT 7  
 G56978  
 collagen alpha 1(I) chain - bovine (fragment)  
 N:Alternate names: collagen alpha 3(XI) chain  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 03-Oct-1995 #sequence\_revision 03-Oct-1995 #text\_change 03-Oct-1995  
 C:Accession: G56978  
 R:Wu, J.C.; Eyre, D.R.  
 J. Biol. Chem. 270, 18865-18870, 1995  
 A:Title: Structural analysis of cross-linking domains in cartilage type XI collagen. In  
 A:Reference number: A56978; MUID:95370194; PMID:7642541  
 A:Accession: G56978  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-9 <WU>  
 A:Note: the residue designated 'X' is modified lysine in collagen 3(XI) some cross-link

Query Match 32.7%; Score 17; DB 2; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 EYGP 8  
 ||||  
 DB 4 EXGP 7

RESULT 8  
 PC2171  
 triacylglycerol lipase (EC 3.1.1.3) I A chain - Rhizopus niveus (strain IFO 4759) (frag  
 C:Species: Rhizopus niveus  
 C:Date: 03-May-1994 #sequence\_revision 07-Oct-1994 #text\_change 09-Jul-2004  
 C:Accession: PC2171  
 R:Kohno, M.; Kugimiyu, W.; Hashimoto, Y.; Morita, Y.  
 Biosci. Biotechnol. Biochem. 58, 1007-1012, 1994  
 A:Title: Purification, characterization, and crystallization of two types of lipase from  
 A:Reference number: PC2171; MUID:94319059; PMID:7765029  
 A:Accession: PC2171  
 A:Molecule type: protein  
 A:Residues: 1-10 <KOH>  
 A:Cross-references: UNIPROT:O74166  
 C:Comment: This enzyme catalyzes the hydrolysis of the ester bonds.  
 C:Keywords: carboxylic ester hydrolase

Query Match 32.7%; Score 17; DB 2; Length 10;  
 Best Local Similarity 50.0%; Pred. No. 3.4e+03;  
 Matches 5; Conservative 1; Mismatches 0; Indels 4; Gaps 1;

OY 1 ALPIEXGPLV 10  
 |||||

Db 4 ALP----PLI 9

## RESULT 9

IG kappa chain V-II region (SD1) - human (fragment)  
 C/Species: Homo sapiens (man)  
 C/Date: 23-Oct-1992 #sequence\_revision 23-Oct-1992 #text\_change 16-Aug-1996  
 C/Accession: U0520  
 R/Anker, R.; Conley, M.E.; Pollok, B.A.  
 J. Exp. Med. 169, 2109-2119, 1989  
 A/Title: Clonal diversity in the B cell repertoire of patients with X-linked agammaglobulinemia  
 A/Reference number: U0511, PMID:8927957, PMID:2786547  
 A/Accession: U0520  
 A/Molecule type: mRNA  
 A/Residues: 1-5 <ANK>  
 A/Note: the sequence shown here is one of eight productive V-D-J mu chain rearrangements  
 A/Note: a stop codon terminates the sequence in the V region  
 C/Keywords: heterotetramer; immunoglobulin  
 F:1-5/Domain: V kappa region <VRB>

Query Match 30.8%; Score 16; DB 2; Length 5;  
 Best Local Similarity 75.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 EYGP 8  
 : : :  
 : : :  
 Db 2 EGGP 5

## RESULT 10

DNA topoisomerase II alpha - Chinese hamster (fragment)  
 C/Species: Cricetus griseus (Chinese hamster)  
 C/Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999  
 C/Accession: I48086  
 R/Ng, S.W.; Eder, J.P.; Schnipper, L.E.; Chan, V.T.W.  
 J. Biol. Chem. 270, 25850-25858, 1995  
 A/Title: Molecular cloning and characterization of the promoter for the Chinese hamster  
 A/Reference number: I48086, PMID:96029684; PMID:7592770  
 A/Accession: I48086  
 A/Status: Preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-7 <RES>  
 A/Cross-references: EMBL:U34196; NID:G1041231; PIDN:AC52315.1; PID:G1041232

Query Match 30.8%; Score 16; DB 2; Length 7;  
 Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 IEYGP 9  
 : : :  
 : : :  
 Db 1 MELSP 6

## RESULT 11

adipokinetic hormone - cockchafer  
 C/Species: Melolontha melolontha (cockchafer)  
 C/Date: 19-Mar-1997 #sequence\_revision 31-Oct-1997 #text\_change 09-Jul-2004  
 C/Accession: S15422  
 R/Gaede, G.  
 Biochem. J. 275, 671-677, 1991  
 A/Title: A unique charged tyrosine-containing member of the adipokinetic hormone/red-pigment concentrating hormone family  
 A/Reference number: S15422, PMID:91248100; PMID:2039445  
 A/Accession: S15422  
 A/Molecule type: protein  
 A/Residues: 1-8 <BIO>  
 A/Cross-references: UNIPROT:P25423  
 C/Superfamily: adipokinetic hormone  
 C/Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid  
 F:1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental  
 F:8/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 30.8%; Score 16; DB 2; Length 8;  
 Best Local Similarity 40.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 IEYGP 8  
 : : :  
 : : :  
 Db 2 LNYGP 6

## RESULT 12

adipokinetic hormone - dor beetle  
 C/Species: Geotrupes stercorarius (dor beetle)  
 C/Date: 28-Oct-1997 #sequence\_revision 31-Oct-1997 #text\_change 09-Jul-2004  
 C/Accession: A58641  
 R/Gaede, G.  
 Biochem. J. 275, 671-677, 1991  
 A/Title: A unique charged tyrosine-containing member of the adipokinetic hormone/red-pigment concentrating hormone family  
 A/Reference number: S15422; PMID:91248100; PMID:2039445  
 A/Accession: A58641  
 A/Molecule type: protein  
 A/Residues: 1-8 <BIO>  
 A/Cross-references: UNIPROT:P25423  
 C/Superfamily: adipokinetic hormone  
 C/Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid  
 F:1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental  
 F:8/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 30.8%; Score 16; DB 2; Length 8;  
 Best Local Similarity 40.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 IEYGP 8  
 : : :  
 : : :  
 Db 2 LNYGP 6

## RESULT 13

IG gamma chain C region (gamma-1) - human (fragment)  
 C/Species: Homo sapiens (man)  
 C/Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Aug-1996  
 C/Accession: P0368  
 R/Millili, M.; Fougereau, M.; Guglielmi, P.; Schiff, C.  
 Mol. Immunol. 28, 753-761, 1991  
 A/Title: Early occurrence of immunoglobulin isotype switching in human fetal liver.  
 A/Reference number: P0368; PMID:9132348; PMID:1906381  
 A/Accession: P0368  
 A/Molecule type: mRNA  
 A/Residues: 1-8 <MIL>  
 A/Experimental source: fetal liver  
 C/Keywords: immunoglobulin

Query Match 30.8%; Score 16; DB 2; Length 8;  
 Best Local Similarity 33.3%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 IEYGP 9  
 : : :  
 : : :  
 Db 1 LHOGEI 6

## RESULT 14

neuropeptide - flower beetle (Pachnoda marginata)  
 C/Species: Pachnoda marginata  
 C/Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
 C/Accession: S21663  
 R/Gaede, G.; Lopata, A.; Kellner, R.; Rinehart, K.L.  
 Biol. Chem. Hoppe-Seyler 373, 133-142, 1992  
 A/Title: Primary structures of neuropeptides isolated from the corpora cardiaca of various insects

A:Reference number: S21663; MUID:92265187; PMID:1586453  
A:Accession: S21663  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 18 <GAE>  
A:Cross-references: UNIPROT:P25423

Query Match                      30.8%; Score 16; DB 2; Length 8;  
Best Local Similarity      40.0%; Pred. No. 2.8e+05;  
Matches      2; Conservative      1; Mismatches      2; Indels      0; Gaps      0;

QY                      4 IEVGP 8  
                              | | |  
DB                      2 LNYSP 6

## RESULT 15

S65388  
cytochrome-c oxidase (EC 1.9.3.1) chain VII c, hepatic - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S65388; S65389  
R:Schlegger, H.; Noack, H.; Halangk, W.; Brandt, U.; von Jagow, G.  
Eur. J. Biochem. 230, 235-241, 1995  
A:Title: Cytochrome-c oxidase in developing rat heart. Enzymic properties and amino-term  
A:Reference number: S65372; MUID:95324529; PMID:7601105  
A:Accession: S65388  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-10 <SCH>  
A:Cross-references: UNIPROT:P80432  
A:Accession: S65389  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-10 <SC2>  
C:Superfamily: cytochrome-c oxidase chain VIIc  
C:Keywords: oxidoreductase

Query Match                      30.8%; Score 16; DB 2; Length 10;  
Best Local Similarity      75.0%; Pred. No. 5.4e+03;  
Matches      3; Conservative      0; Mismatches      1; Indels      0; Gaps      0;

QY                      5 EYGP 8  
                              | | |  
DB                      4 BEGP 7

Search completed: November 30, 2004, 08:14:27  
Job time : 12.6667 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:40:12 / Search time 78.333 Seconds  
(without alignments)  
73.452 Million cell updates/sec

Title: US-10-008-377a-6

Perfect score: 52

Sequence: 1 ALPIEYGLV 10

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 2971

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: UniProt\_02.\*

1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	23.5	45.2	10 2 P83154	P83154 anabaena sp
2	20	38.5	9 2 Q7RSP2	Q7RSP2 plasmidum
3	19	36.5	8 1 ALI6 CARMA	Q977H9 haemophilus
4	19	36.5	8 1 ALI6 CARMA	P81819 carcinus ma
5	18	34.6	8 2 Q16468	Q16468 homo sapien
6	17	32.7	7 1 MNPI_LBPDE	Q83332 murine hepa
7	17	32.7	8 1 ALI4_CYPPO	P42984 leptinotars
8	17	32.7	8 1 ALI4_CYPPO	P42984 leptinotars
9	17	32.7	8 2 Q7KXV5	P81815 cydia pomon
10	17	32.7	8 2 Q7KXV5	Q7KXV5 homo sapien
11	17	32.7	9 2 Q9X3X1	Q9X3X1 prochloroco
12	17	32.7	9 2 Q9X3X1	Q9X3X1 anthoplaura
13	17	32.7	10 2 Q6JVP0	Q6JVP0 welliaetonia
14	17	32.7	10 2 Q6JVP0	Q6JVP0 cecopaprus e
15	17	32.7	10 2 Q6JVP6	Q6JVP6 elaphandra
16	17	32.7	10 2 Q6JVP8	Q6JVP8 dimerosetum
17	17	32.7	10 2 Q6JVP0	Q6JVP0 dimerosetum
18	17	32.7	10 2 Q6JVP0	Q6JVP0 baillimora x
19	17	32.7	10 2 Q6JVP0	Q6JVP0 angelphytum
20	17	32.7	10 2 Q6JVP0	Q6JVP0 angelphytum
21	17	32.7	10 2 Q6JVP0	Q6JVP0 angelphytum
22	17	32.7	10 2 Q6JVP0	Q6JVP0 angelphytum
23	17	32.7	10 2 Q6JVP0	Q6JVP0 angelphytum
24	17	32.7	10 2 Q6JVP0	Q6JVP0 angelphytum
25	17	32.7	10 2 Q6JVP0	Q6JVP0 angelphytum
26	17	32.7	10 2 Q6JVP0	Q6JVP0 angelphytum
27	17	32.7	10 2 Q6JVP0	Q6JVP0 angelphytum
28	17	32.7	10 2 Q6JVP0	Q6JVP0 angelphytum
29	17	32.7	10 2 Q6JVP0	Q6JVP0 angelphytum
30	17	32.7	10 2 Q6JVP0	Q6JVP0 angelphytum
31	17	32.7	10 2 Q6JVP0	Q6JVP0 angelphytum

32	17	32.7	10 2 AAR14768	AAR14768 cecopaprus
33	17	32.7	10 2 AAR14772	AAR14772 welliaetonia
34	16	30.8	5 1 FARP_CHICK	P83308 gallus gall
35	16	30.8	6 1 SAPP_SEPOF	P83569 sepiia offic
36	16	30.8	8 1 AKH_MELMU	P25423 meliolonthe
37	16	30.8	8 1 ALI7 CARMA	P81809 carcinus ma
38	16	30.8	8 2 Q7OKG9	Q7OKG9 sus scrofa
39	16	30.8	8 2 P82598	P82598 rattus norv
40	16	30.8	8 2 CAE06499	CAE06499 sus scrof
41	16	30.8	8 2 CAE06500	CAE06500 sus scrof
42	16	30.8	8 2 CAE06501	CAE06501 sus scrof
43	16	30.8	8 2 CAE06502	CAE06502 sus scrof
44	16	30.8	9 2 Q9UE09	Q9UE09 homo sapien
45	16	30.8	9 2 P82568	P82568 streptococc

#### ALIGNMENTS

RESULT 1				
ID	P83154	PRELIMINARY;	PRT;	10 AA.
AC	P83154			
DT	01-MAR-2002 (TREMURel. 20, Created)			
DT	01-MAR-2002 (TREMURel. 20, Last sequence update)			
DE	01-OCT-2003 (TREMURel. 25, Last annotation update)			
DE	Phycobillosome rod-core linker polypeptide cpoc3 (Fragment).			
OS	Anabaena sp. (strain 131)			
OC	Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.			
OK	NCBI_TaxID=29412;			
RN	(1)			
RA	SEQUENCE.			
RA	Apte S.K., Uhlemann E., Schmid R., Alenddorf K.,			
RL	Submitted (OCT-2001) to Swiss-Prot.			
CC	-1- FUNCTION: Rod-core linker protein required for attachment of			
CC	phycocyanin to allophycocyanin in cores of phycobillosomes.			
CC	-1- FUNCTION: Linker polypeptides determine the state of aggregation			
CC	and the location of the disk-shaped phycobilliprotein units within			
CC	the phycobillosome and modulate their spectroscopic properties in			
CC	order to mediate a directed and optimal energy transfer.			
CC	-1- SUBUNIT: The phycobillosome is a hemidiscoidal structure that is			
CC	composed of two distinct substructures: a core complex (that			
CC	contains phycobilliproteins) and a number of rods radiating from			
CC	the core.			
CC	-1- SUBCELLULAR LOCATION: Thylakoid membrane.			
CC	-1- SIMILARITY: TO OTHER PHYCOBILLSOME LINKER PROTEINS.			
DR	GO; GO:0016020; C:membrane; IEA.			
DR	GO; GO:0030899; C:phycobillosome; IEA.			
DR	GO; GO:0009579; C:thylakoid; IEA.			
DR	GO; GO:0015979; P:photosynthesis; IEA.			
KW	Membrane; Photosynthesis; Phycobillosome; Thylakoid.			
FT	NON TER			
FT	SEQUENCE 10 AA; 1144 MW; 2F9B662B5B172737 CRC64;			
QY	Query Match	45.2%; Score 23.5; DB 2; Length 10;		
QY	Best Local Similarity	66.7%; Pred. No. 1.6e+03;		
DB	Matches	6; Conservative 1; Mismatches 1; Indels 1; Gaps 1;		
DB	1 ALP-IEYGP 8			
DB	1 ALPIEYXP 9			
RESULT 2				
ID	Q7RSP2	PRELIMINARY;	PRT;	9 AA.
AC	Q7RSP2			
DT	01-MAR-2004 (TREMURel. 26, Created)			
DT	01-MAR-2004 (TREMURel. 26, Last sequence update)			
DT	01-MAR-2004 (TREMURel. 26, Last annotation update)			
DE	Hypothetical protein.			
GN	Name=PY00313;			
OS	Plasmidium yoeili yoeili.			

```

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RN [1]_Taxid=73239;
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=1238865;
RA Carlton J.M., Anguilo S.V., Suh B.B., Kocif T.W., Petrea M.,
RA Silva J.C., Ermolaeva W.D., Allen J.E., Selenyut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.P., Bidwell S.L.,
RA Shalom S.J., van Aken S.E., Riedmuller S.B., Feldblum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoabli A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Valdey A.B.,
RA van Lin L.H., Jarne C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carnocci D.O.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii."
RL Nature 419:512-519(2002).
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABL000088; EAA22678.1; -.
KM Hypothetical protein.
SQ SEQUENCE 9 AA; 1067 MW; 054CBA46D0505EB0 CRC64;

QY 4 IEVGP 8
DB 5 IGVSP 9

Query Match 38.5%; Score 20; DB 2; Length 9;
Best Local Similarity 60.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

RESULT 3
Q5R7H9 PRELIMINARY; PRT; 9 AA.
ID Q5R7H9;
AC Q5R7H9;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Lipoprotein (Fragment).
GN Name=lpd;
OS Haemophilus influenzae;
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_Taxid=727;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=98083063; PubMed=9422600;
RA Martin K., Morlin G., Smith A., Nordyke A., Eisenstark A., Golomb M.,
RT "The tryptophanase gene cluster of Haemophilus influenzae type D:
RT evidence for horizontal gene transfer."
RL J. Bacteriol. 180:107-118(1998).
DR EMBL; AF003252; AAB96582.1; -.
FT NON_TER
SQ SEQUENCE 9 AA; 1152 MW; 35A017673B4412D7 CRC64;

Query Match 38.5%; Score 20; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 PTEYGP 8
DB 2 PVRYLP 7

RESULT 4
AL16 CAPMA STANDARD; PRT; 8 AA.
ID AL16 CAPMA
AC P81819;
DT 30-MAY-2000 (rel. 39, Created)

```

```

DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Carcinustatin 15.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Decapoda; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_Taxid=6759;
RN [1]_
RP TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duvé H., Johnsen A.H., Maestro J.-L., Scott A.G., Jares P.P.,
RA Thøpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allostatin superfamily in the shore crab Carcinus maenas."
RL Eur. J. Biochem. 250:727-734(1997).
CC -1- FUNCTION: May act as a neurotransmitter or neuromodulator.
CC -1- SIMILARITY: Belongs to the allostatin family.
CC AMIATION: Direct protein sequencing; Multigene family; Neuropeptide.
KM MOD RES 8 Leucine amide.
FT MOD RES 8
SQ SEQUENCE 8 AA; 913 MW; 7C286B45A8476878 CRC64;

Query Match 36.5%; Score 19; DB 1; Length 8;
Best Local Similarity 60.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 PTEYGP 7
DB 3 PVSYP 7

RESULT 5
O16468 PRELIMINARY; PRT; 8 AA.
ID O16468;
AC O16468;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE H.sapiens DNA for cosmid cC13-1134 PCR primer 1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=96435920; PubMed=8838806;
RA James L.A., Ogilvie D.J., Yamakawa K., Nakamura Y., Stirling C.J.,
RA Anand R.;
RT "Walking, cloning and mapping with YACs in 3q27. Localisation of 5
RT ESTs including 3 members of the Cystatin gene family and
RT identification of CG islands."
RL Genomics 32:425-430(1996).
DR EMBL; X88976; CAA61407.1; -.
FT NON_TER
SQ SEQUENCE 8 AA; 925 MW; FD5411A7376871E6 CRC64;

Query Match 36.5%; Score 19; DB 2; Length 8;
Best Local Similarity 75.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 YGPL 9
DB 2 HGPL 5

RESULT 6
O83332 PRELIMINARY; PRT; 8 AA.
ID O83332;
AC O83332;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)

```

DE HE (Fragment).  
 OS Murine hepatitis virus.  
 OC Virusess; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
 CC Coronaviridae; Coronavirinae; Group 2 species.  
 CX NCBI\_TaxID=11138;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JHM;  
 RX MEDLINE=95191005; PubMed=7884877;  
 RA Kim K.H., Makino S.;  
 RT "Two murine coronavirus genes suffice for viral RNA synthesis.";  
 RL J. Virol. 69:2313-2321(1995).  
 DR EMBL: U19933; AA69002.1; -.  
 FT NON\_TER 1 1  
 FT MOD\_RRS 8 8  
 FT SEQUENCE 8 AA; 907 MW; 922735B1735A2CD CRC64;

Query Match 34.6%; Score 18; DB 2; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Cy 1 ALPIRY 6  
 Db 1 AYSLEY 6

RESULT 7  
 NMPL LEDE STANDARD; PRT; 7 AA.  
 ID NMPL LEDE  
 AC P42984;  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Myotropic neuropilin-1 (led-nmp-1)  
 OS Lepidoptera: decemlineata (Colorado potato beetle).  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neuroptera; Megaloptera; Coleoptera; Polyphaga; Cuculiformia;  
 CC Phycophaga; Chrysomelidae; Chrysomelinae;  
 CC Chrysomelinae; Leptinotarsa.  
 CC NCBI\_TaxID=7539;  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RC TISSUE=Head;  
 RX MEDLINE=95380343; PubMed=7651886;  
 RA Spittels K., Vankeerberghen A., Schoofs L., Torckens S.,  
 RA Grauwels L., van Leuven F., de Loof A.;  
 RT "Identification, characterization, and immunological localization of a  
 RT novel myotropic neuropilin in the Colorado potato beetle,  
 RT Leptinotarsa decemlineata.";  
 RL Peptides 16:365-374(1995).  
 CC -1- FUNCTION: Myotropic peptide. Stimulates the contractions of the  
 CC oviduct.  
 CC Annotation: Direct protein sequencing; Neuropeptide.  
 KW MOD\_RRS 7 7  
 KW MOD\_RRS 7 7  
 KW SEQUENCE 7 AA; 705 MW; 6DD373768745B5DB0 CRC64;

Query Match 32.7%; Score 17; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 7 GPL 9  
 Db 4 GPL 6

RESULT 8  
 ALM4 CYDPO STANDARD; PRT; 8 AA.  
 ID ALM4 CYDPO  
 AC P82155;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DS Cydactactin 4.

OS Cydia pomonella (Codling moth).  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neuroptera; Megaloptera; Lepidoptera; Glossata; Ditrysia;  
 CC Tortricidae; Tortricidae; Olethreutinae; Cydia.  
 CX NCBI\_TaxID=82600;  
 RN [1]  
 RP SEQUENCE  
 RC TISSUE=Larva;  
 RX MEDLINE=98054539; PubMed=93928239;  
 RA Duvé H., Johnson A.H., Maestri U.-L., Scott A.G., Minstaneley D.,  
 RA Davey M., East P.D., Thorpe A.;  
 RT "Lepidopteran peptides of the allatostatin superfamily.";  
 RL Peptides 18:1301-1309(1997).  
 CC -1- SIMILARITY: Belongs to the allatostatin family.  
 CC Annotation: Direct protein sequencing; Neuropeptide.  
 KW MOD\_RRS 8 8  
 KW MOD\_RRS 8 8  
 KW SEQUENCE 8 AA; 910 MW; 922879D5A547740D CRC64;

Query Match 32.7%; Score 17; DB 1; Length 8;  
 Best Local Similarity 42.9%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Cy 1 ALPIRY 7  
 Db 1 ARYSTRG 7

RESULT 9  
 ALMS CALVO STANDARD; PRT; 8 AA.  
 ID ALMS CALVO  
 AC P41841;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Callaostatin 1 (Hyp3)Met-callaostatin 1)  
 OS Calliphora vomitoria (Blue blowfly).  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neuroptera; Megaloptera; Diptera; Brachycera; Muscomorpha; Oestroidea;  
 CC Calliphoridae; Calliphora.  
 CC NCBI\_TaxID=27454;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Head;  
 RX MEDLINE=93211980; PubMed=8460157;  
 RA Duvé H., Johnson A.H., Scott A.G., Yu C.G., Yagi K.J., Tobe S.S.,  
 RA Thorpe A.;  
 RT "Callaostatin: neuropeptides from the blowfly Calliphora vomitoria  
 RT with sequence homology to cockroach allatostatin.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:2456-2460(1993).  
 RN [2]  
 RP CHARACTERIZATION, AND HYDROXYLATION.  
 RC TISSUE=Head;  
 RX MEDLINE=9434269; PubMed=8063725;  
 RA Duvé H., Johnson A.H., Scott A.G., East P., Thorpe A.;  
 RT "[Hyp3]Met-callaostatin. Identification and biological properties of  
 RT a novel neuropeptide from the blowfly Calliphora vomitoria.";  
 RL J. Biol. Chem. 269:21059-21066(1994).  
 CC -1- FUNCTION: May act as a neurotransmitter or neuromodulator and play  
 CC a role in the integration of information within the brain. May be  
 CC involved in the control of visceral muscles due to its ability to  
 CC behave as potent inhibitors of peristaltic movements. May also  
 CC fulfill a neurohormonal role on muscles of the gut and heart.  
 CC -1- TISSUE SPECIFICITY: Neurons within brain and abdominal ganglion.  
 CC -1- SIMILARITY: Belongs to the allatostatin family.  
 DR PIR: E47393; E47393.  
 KW Annotation: Direct protein sequencing; Hydroxylation; Neuropeptide.  
 KW MOD\_RRS 3 3  
 KW MOD\_RRS 3 3  
 KW SEQUENCE 8 AA; 883 MW; 7D9679CAB5477768 CRC64;

Query Match 32.7%; Score 17; DB 1; Length 8;  
 Best Local Similarity 40.0%; Pred. No. 1.8e+06;  
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 P1EY6 7  
DB 3 PYDF6 7

## RESULT 10

O7KVV5 PRELIMINARY; PRT; 8 AA.  
AC 07KVV5;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE Tenascin-X (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97081760; PubMed=8923003;  
RA Speck W., Barry F., Miller W.L.;  
RT "Alternate promoters and alternate splicing of human tenascin-X, a  
RT gene with 5' and 3' ends buried in other genes."  
RL Hum. Mol. Genet. 5:1749-1758(1996).  
DR EMBL; US2700; AAC50882.1;  
FT NON TER  
SQ SEQUENCE 8 AA; 924 MW; FE2DCB46DDC76696 CRC64;

Query Match 32.7%; Score 17; DB 2; Length 8;  
Best Local Similarity 40.0%; Pred. No. 1.8e+06;  
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 LP1EY 6  
DB 2 MPAGY 6

## RESULT 11

O9X3K1 PRELIMINARY; PRT; 8 AA.  
AC O9X3K1;  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
DE Cytochrome b (Fragment).  
GN Name=petB;  
OS Prochlorococcus sp.  
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;  
OC Prochlorococcus.  
OX NCBI\_TaxID=1220;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Urbach E., Chisholm S.W.;  
RT "Genetic diversity in Prochlorococcus populations flow cytometrically  
RT sorted from the Sargasso Sea and Gulf Stream."  
RL Limnol. Oceanogr. 43:1615-1630(1998).  
DR EMBL; AF070193; AAD33233.1;  
FT NON TER  
SQ SEQUENCE 8 AA; 799 MW; 10376865B72866D3 CRC64;

Query Match 32.7%; Score 17; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GPL 9  
DB 6 GPL 8

## RESULT 12

O9TWV0 PRELIMINARY; PRT; 9 AA.  
AC O9TWV0;

DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
DE Antho-RPAMIDE-NEUROPEPTIDE.

OS Anthopleura elegantissima (Sea anemone).  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actinaria;  
OC Nymanthea; Actinidae; Anthopleura.  
OX NCBI\_TaxID=6110;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93126143; PubMed=1480510;  
RA Carstensen K., Rinehart K.L., McFarlane I.D., Grimmelikhuijzen C.J.;  
RT "Isolation of Leu-Pro-Gly-Pro-Leu-Pro-Arg-Pro-NH<sub>2</sub> (Antho-RPamide),  
RT an N-terminally protected, biologically active neuropeptide from sea  
RT anemones."  
RL Peptides 13:851-857(1992).  
SQ SEQUENCE 9 AA; 943 MW; 2908176737686777 CRC64;

Query Match 32.7%; Score 17; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GPL 9  
DB 4 GPL 6

## RESULT 13

O6UVPO PRELIMINARY; PRT; 10 AA.  
AC O6UVPO;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE PetB (Fragment).  
GN Name=petB;  
OS Wollastonia biflora (Beach sunflower) (Wedelia biflora).  
OC Chloroplast  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae;  
OX NCBI\_TaxID=101473;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Dias de Moraes M., Panero J.L., Semir J.;  
RT Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY297619; AAR14772.1;  
DR GO; GO:0009507; Chloroplast; IEA.  
KM Chloroplast.  
FT NON TER  
SQ SEQUENCE 10 AA; 1068 MW; 57F74565B05866D3 CRC64;

Query Match 32.7%; Score 17; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 2.9e+04;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GPL 9  
DB 8 GPL 10

## RESULT 14

O6UVP3 PRELIMINARY; PRT; 10 AA.  
AC O6UVP3;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE PetB (Fragment).  
GN Name=petB;  
OS Otopappus epaleaceus.  
OG Chloroplast.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae;  
 OC Otocarpus.  
 OK NCBI\_TaxID=243770;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Dias de Moraes M., Panero J.L., Semir J.  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY297615; ARI14764.1; -  
 DR GO:0009507; C:chloroplast; IEA.  
 KM Chloroplast.  
 FT NON TER 1  
 SQ SEQUENCE 10 AA; 1068 MW; 57F74565B05866D3 CRC64;

Query Match 32.7%; Score 17; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+04;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GPL 9  
 DB 8 GPL 10

RESULT 15  
 O6JVP6 PRELIMINARY; PRT; 10 AA.  
 AC O6JVP6;  
 DT 05-JUL-2004 (TREMURel. 27, Created)  
 DT 05-JUL-2004 (TREMURel. 27, Last sequence update)  
 DE 05-JUL-2004 (TREMURel. 27, Last annotation update)  
 DE PEB (Fragment).  
 GN Name=peb;  
 OS Elaphandra paucipunctata.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae;  
 OC Elaphandra  
 OK NCBI\_TaxID=243766;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Dias de Moraes M., Panero J.L., Semir J.  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY297615; ARI14764.1; -  
 DR GO:0009507; C:chloroplast; IEA.  
 KM Chloroplast.  
 FT NON TER 1  
 SQ SEQUENCE 10 AA; 1068 MW; 57F74565B05866D3 CRC64;

Query Match 32.7%; Score 17; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+04;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GPL 9  
 DB 8 GPL 10

Search completed: November 30, 2004, 08:13:42  
 Job time : 80.333 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:34:52 ; Search time 55.3333 Seconds  
(without alignments)  
64,831 Million cell updates/sec

Title: US-10-008-377A-6  
Perfect score: 52  
Sequence: 1 ALPIEXGPLV 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 444336

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: geneseqp23sep04:\*  
2: geneseqp19sep04:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003ds:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	10	6	AAO22988 Human p45
2	34	65.4	8	2	AAW31323 Dictyocaul
3	34	65.4	10	2	AAV04480 cis-epoxy
4	29	55.8	10	6	ADM41407 Human bet
5	28	53.8	10	3	AA809769 IGFBP-2 I
6	27	51.9	8	2	AA81470 Human cor
7	27	51.9	8	4	AA81470 Human cor
8	26	50.0	8	4	AA81470 Human cor
9	26	50.0	8	4	AA81470 Human cor
10	26	50.0	8	4	AA81470 Human cor
11	26	50.0	8	4	AA81470 Human cor
12	26	50.0	8	4	AA81470 Human cor
13	26	50.0	8	4	AA81470 Human cor
14	26	50.0	8	4	AA81470 Human cor
15	26	50.0	8	4	AA81470 Human cor
16	26	50.0	8	4	AA81470 Human cor
17	26	50.0	8	4	AA81470 Human cor
18	26	50.0	8	4	AA81470 Human cor
19	26	50.0	8	4	AA81470 Human cor
20	26	50.0	8	4	AA81470 Human cor
21	26	50.0	8	4	AA81470 Human cor
22	26	50.0	8	4	AA81470 Human cor
23	26	50.0	8	4	AA81470 Human cor
24	26	50.0	8	4	AA81470 Human cor
25	26	50.0	8	4	AA81470 Human cor

26	26	50.0	9	6	ABJ60273	ABJ60273 184P1E2-x
27	26	50.0	9	6	ABJ58282	ABJ58282 184P1E2-x
28	26	50.0	9	6	ABJ58507	ABJ58507 184P1E2-x
29	26	50.0	9	6	ABJ61111	ABJ61111 184P1E2-x
30	26	50.0	10	2	AAW70194	AAW70194 N-termina
31	26	50.0	10	4	AA847586	AA847586 Ag85 comp
32	26	50.0	10	4	AA805660	AA805660 N-termina
33	26	50.0	10	4	AA805651	AA805651 N-termina
34	26	50.0	10	4	AA805656	AA805656 N-termina
35	26	50.0	10	6	ABJ58230	ABJ58230 184P1E2-x
36	26	50.0	10	6	ABJ67866	ABJ67866 184P1E2-x
37	26	50.0	10	6	ABJ69133	ABJ69133 184P1E2-x
38	26	50.0	10	6	ABJ57712	ABJ57712 184P1E2-x
39	26	50.0	10	6	ABJ58481	ABJ58481 184P1E2-x
40	26	50.0	10	6	ABJ66762	ABJ66762 184P1E2-x
41	26	50.0	10	6	ABJ68412	ABJ68412 184P1E2-x
42	26	50.0	10	6	ABJ58638	ABJ58638 184P1E2-x
43	26	50.0	10	6	ABJ68202	ABJ68202 184P1E2-x
44	26	50.0	10	6	ABJ68552	ABJ68552 184P1E2-x
45	26	50.0	10	6	ABJ69211	ABJ69211 184P1E2-x

## ALIGNMENTS

RESULT 1  
AAO22988  
ID AAO22988 standard; peptide; 10 AA.  
XX  
AC AAO22988;  
XX  
DT 17-SEP-2003 (first entry)  
XX  
DE Human p45-65 ALK-derived HLA-A\*0201 restricted CTL epitope peptide.  
XX  
KW HLA-A\*0201 binding; anaplastic lymphoma kinase; ALK; CTL epitope; ALCL;  
XX cytotoxic T-lymphocyte; receptor tyrosine; NPW/ALK fusion protein; human;  
XX oncogenic; anaplastic large cell lymphoma; neuroblastoma; gene therapy;  
XX t(2;5) translocation; neoplasia; cytostatic; immunotherapy; alkoma;  
KW p45-65.  
XX  
OS Homo sapiens.  
XX  
PN WO2003042243-A2.  
XX  
PD 22-MAY-2003.  
XX  
PF 14-NOV-2002; 2002WO-EP012764.  
XX  
XX 15-NOV-2001; 2001US-00008377.  
XX  
PA (MAST-) 1ST NAZ STUDIO E CURA DEI TUMORI.  
XX (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
XX  
PI Gambacorti-Passerini C, Passoni L;  
XX  
DR MPI; 2003-441791/41.  
XX  
PT New HLA-A\*0201-binding Anaplastic Lymphoma Kinase (ALK) peptide,  
XX useful for preparing a composition for treating ALK-positive lymphoma,  
XX neuroblastoma or ALK-expressing neoplasia.  
XX  
XX Claim 1; Page 7; 33pp; English.  
XX  
XX The invention relates to a novel HLA-A\*0201-binding anaplastic lymphoma  
XX kinase (ALK)-derived peptide, of 9-10 amino acids, which may act as a  
XX cytotoxic T-lymphocyte (CTL)-epitope. ALK is a receptor tyrosine kinase  
XX which forms part of the NPW/ALK fusion protein (oncogenic lymphoma kinase  
XX fusion protein) in ALCL (anaplastic large cell lymphoma)-derived cell  
XX lines. More than 50% of ALCL cases possess a t(2;5) chromosomal  
XX translocation that leads to the expression of the NPW/ALK fusion protein  
XX which forms a potent oncogene when constitutively activated. Translocated  
XX ALK is a widely expressed tumour-associated antigen characteristic of ALK

CC -positive lymphomas, neuroblastomas and ALK-expressing neoplasias. The  
 CC peptide of the invention demonstrates cytostatic activity and induce an  
 CC MHC (major histocompatibility complex) class I restricted cytotoxic  
 CC lymphocyte response against tumour cells expressing the NPW/ALK fusion  
 CC protein. Hence, the peptides may be utilised during the treatment, via  
 CC immunotherapy, of ALK-positive lymphomas such as alkomas, neuroblastomas  
 CC and ALK-expressing neoplasias. In addition, the peptides may be used  
 CC during gene therapy. The current sequence is that of the human p456-65  
 CC ALK-derived HLA-A\*0201 restricted CTL epitope peptide of the invention  
 XX

SO Sequence 10 AA;

Query Match 100.0%; Score 52; DB 6; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0064;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPIEYGPLV 10  
 |||||  
 Db 1 ALPIEYGPLV 10

RESULT 2  
 AAW31323  
 ID AAW31323 standard; peptide; 8 AA.

AC AAW31323;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 13-MAR-1998 (first entry)

DE Dictyocaulus viviparus DV18 peptide fragment 3.

KW Immunogenic protein; DV18; lungworm; vaccine; immunity; dictyocauliasis;  
 KW cattle; immunosassay.

OS Dictyocaulus viviparus.

PN EP785253-A1.

PD 23-JUL-1997.

PF 27-DEC-1996; 96EP-00120947.

PR 19-JAN-1996; 96DE-01001754.

PA (FARH) HOECHST AG.

PI Hofmann J, Schmid K;

XX WPI; 1997-365928/34.

PT Immunogenic Dictyocaulus viviparus lungworm protein - for use in vaccines  
 PT and immunosassays.

PS Example 10; Page 6; 17pp; German.

XX This sequence represents a peptide fragment of the immunogenic protein  
 CC DV18 isolated from adult lungworms. The N-terminal of this amino acid can  
 CC be joined to the amino acids Asp, Asn, Ser, Gly or Arg. DV18 can be used  
 CC in vaccines for immunising cattle against dictyocauliasis and in an ELISA  
 CC immunosassay for determining DV18-specific antibodies in the blood of  
 CC cattle. (Updated on 25-MAR-2003 to correct PR field.)  
 XX

SO Sequence 8 AA;

Query Match 65.4%; Score 34; DB 2; Length 8;  
 Best Local Similarity 85.7%; Pred. No. 1.7e+06;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LPIEYGP 8  
 |||||  
 Db 2 LPIEYNP 8

RESULT 3  
 AAY04480  
 ID AAY04480 standard; peptide; 10 AA.

AC AAY04480;

DT 30-JUN-1999 (first entry)

DE cis-epoxysuccinate hydrolase internal peptide #2.

KW Rhodococcus rhodochrous; cis-epoxysuccinate hydrolase; microbial;  
 KW epoxide hydrolase; L(+); tartaric acid; food industry; additive;  
 KW soft drink; food preservative; emulsifier.

OS Rhodococcus rhodochrous.

PN EP911392-A1.

PD 28-APR-1999.

PF 24-OCT-1997; 97EP-00870168.

PR 24-OCT-1997; 97EP-00870168.

PA (PURA-) PURATOS NV.

PI Davrin T, Deslee P;

DR WPI; 1999-246412/21.

PT New epoxide hydrolase from Rhodococcus rhodochrous useful for the  
 PT production of L(+); tartaric acid.

PS Example 2; Page 7; 21pp; English.

XX The present sequence is a peptide from Rhodococcus rhodochrous IMGP-18079  
 CC cis-epoxysuccinate hydrolase (epoxide hydrolase). Epoxides are used as  
 CC chiral building blocks in the organic synthesis of fine chemicals,  
 CC especially enantiomerically pure compounds. The epoxide hydrolase can be  
 CC used to hydrolyze an epoxide, especially cis-epoxysuccinate, leading to  
 CC the production of L(+) tartaric acid. This tartaric acid is used in the  
 CC food industry as an additive in soft drinks, a food preservative and a  
 CC raw material for the synthesis of emulsifiers. L(+) tartaric acid for use  
 CC in the food industry can be produced from maleic acid but a mixture of  
 CC both the L(+) and the D(+) forms results and the D(+) form is considered  
 CC harmful to humans. L(+) tartaric acid can also be produced as a by-  
 CC product of wine fermentation but the supply is variable and dependent on  
 CC the climate from year to year. However epoxide hydrolases can be used to  
 CC produce pure L(+) tartaric acid which can therefore be used in the food  
 CC industry. Epoxide hydrolases have been identified in mammals, plants and  
 CC insects but only low levels can be obtained from these sources so are not  
 CC useful for large-scale processes. However, production of enzymes from  
 CC microorganisms culture as claimed for the new enzyme can be applied to  
 CC industrial-scale production  
 XX

SO Sequence 10 AA;

Query Match 65.4%; Score 34; DB 2; Length 10;  
 Best Local Similarity 83.3%; Pred. No. 15;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 PLEIYGP 8  
 |||||  
 Db 2 PLEIYGP 7

RESULT 4  
 ADM41407  
 ID ADM41407 standard; peptide; 10 AA.

AC ADM41407;

XX

DT	03-JUN-2004 (first entry)
XX	
DE	Human beta3 adrenergic receptor loop 2-derived peptide R013JL103.
KM	Beta3 adrenergic receptor; G-protein coupled receptor; receptor;
KM	angiogenesis; antiarteriosclerotic; vulnerable; hypotensive; cardiac;
KM	cerebroprotective; vasotropic; nootropic; neuroprotective; nephroprotective;
KM	cytostatic; antiasthmatic; antidiabetic; immunomodulator;
KM	signal transduction; human.
OS	
OS	Homo sapiens.
XX	
FH	Location/Qualifiers
FT	Key
FT	Modified-site 1 /note= "Myristoylated"
XX	
FN	WO2004022576-A2.
XX	
FN	18-MAR-2004.
XX	
PD	
XX	
PF	03-SEP-2003; 2003WO-US027332.
XX	
PR	03-SEP-2002; 2002US-0407290P.
XX	
PA	(CHILD-) CHILDREN'S HOSPITAL MEDICAL CENT.
PI	
PI	Ben-Sasson S, Reuveni H;
DR	
XX	WPI; 2004-248441/23.
XX	
PT	Identifying candidate compounds for modulating signal transduction
PT	associated with a 7TM receptor comprises identifying a peptide region in
PT	the 7TM receptor (unique region) and synthesizing and testing the
PT	compound.
XX	
PS	Claim 18; SEQ ID NO 9; 77pp; English.
XX	
CC	The present sequence is that of R013JL103, a peptide derived from the loop
CC	2 region ADM41404 of the human beta3 adrenergic receptor, a 7-
CC	transmembrane (7TM) receptor. The invention is based on the discovery
CC	that short peptides from the second loop of 7TM receptors are capable of
CC	selectively modulating signal transduction mediated by that 7TM receptor.
CC	Methods are provided for identifying candidate compounds for the
CC	modulation of signal transduction associated with a 7TM receptor. Such
CC	compounds may comprise the present sequence. They can be used to treat a
CC	disease associated with signal transduction from a 7TM receptor,
CC	including hypertension, stroke, heart failure, neurodegenerative diseases
CC	(including Alzheimer's disease), renal disease, psychiatric disease,
CC	cancer, asthma, diabetes and immune disorders (all claimed). In an
CC	example from the invention, R013JL103 was shown not to increase cAMP
CC	production from beta2-adrenergic transfected HEK293 and B16 cells
CC	compared to the control level, but increased cAMP production 18-fold
CC	after 120 minutes incubation with 37U/L adipocyte cells. It did not
CC	induce melanogenesis from B16 cells.
XX	
XX	
SQ	Sequence 10 AA:
Query Match	55.8%; Score 29; DB 8; Length 10;
Best Local Similarity	62.5%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0	
QY	3 PIRGAPLV 10   :         :
Dd	2 PLRYGALV 9
RESULT 5	
AAB09769	ID AAB09769 standard; peptide; 10 AA.
AC	AAB09769;
DT	01-SEP-2000 (first entry)

XX	IGFBP-2 IGF binding domain derived amino acid sequence SEQ ID NO:138.
XX	
XX	HuMat; insulin like growth factor; IGF; IGFBP; binding domain;
KM	insulin like growth factor binding protein; diabetic complication;
KW	ischemic injury; antagonist; modification; gene therapy; cyostatic;
KW	vasoconstrictor; antidiabetic; antiParkinsonian; neuroprotective; osteopathic;
KW	antiarthritic; vulnery; tranquiliser; neurological disease; head trauma;
KW	Parkinson's disease; amyotrophic lateral sclerosis; multiple sclerosis;
XX	osteoporosis; arthritis.
XX	
XX	Homo sapiens.
XX	
XX	WO200023469-A2.
XX	
XX	27-APR-2000.
XX	
XX	14-OCT-1999; 99NO-USO23839.
XX	
XX	16-OCT-1998; 98US-0104528P.
XX	
XX	(MUSC-) MUSC FOUND RES DEV.
XX	
XX	Rosenzweig SA, Horney MJ;
XX	
XX	WPI; 2000-339652/29.
DR	
PT	New isolated peptide having an insulin-like growth factor domain of an
PT	insulin-like growth factor binding protein, useful for treating or
PT	preventing cancer or diabetic complications, or for treating ischemic
PT	injury.
XX	
PS	Claim 4; Page 14; 106pp; English.
XX	
XX	The present invention describes an isolated peptide (A) comprising an
CC	insulin like growth factor (IGF) binding domain of an IGF-binding protein
CC	(IGFBP) or its modification. (A) binds IGF with at least the same binding
CC	affinity as the full length IGFBP. A peptide from the present invention
CC	can have cyostatic, vasoconstrictor, antidiabetic, antiParkinsonian,
CC	neuroprotective, osteopathic, antiarthritic, vulnery and tranquiliser
CC	activities. The peptide is an IGF inhibitor, IGF antagonist and can be
CC	used in gene therapy. The peptide and antagonists from the present
CC	invention are useful for the treatment or prevention of cancer or
CC	diabetic complications, and for treating ischemic injury. Other diseases
CC	or injuries that can be treated with the fragment or antagonist include
CC	neurologic diseases and injuries, e.g. Parkinson's disease, amyotrophic
CC	lateral sclerosis, head trauma or multiple sclerosis, osteoporosis or
CC	arthritis. The bicyclicated IGF is useful in therapeutic assays for IGFBP
CC	and in screening for IGFBP-mimetics (e.g. IGF antagonists). AAB09616 to
CC	AAB09773 represent amino acid sequences used in the exemplification of
CC	the present invention
SQ	
XX	Sequence 10 AA:
XX	
Query Match	53.8%; Score 28; DB 3; Length 10;
Best Local Similarity	75.0%; Pred. No. 1,9e+02;
Matches	6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY	2 LPIEYGPL 9
DG	1 LPDERGPL 8
RESULT 6	
ID	AAAY81470 standard; peptide; 10 AA.
XX	
AC	AAAY81470;
XX	
DT	03-JUL-2000 (first entry)
XX	
DE	Human corticosteroid precursor N-terminal peptide.

KW Corticosterin precursor; human; Edman degradation; N-terminal;  
 KW polypeptide sequencing apparatus; protein sequence analysis.  
 OS Homo sapiens.  
 XX  
 XX WO200012994-A1.  
 PD 09-MAR-2000.  
 XX  
 XX 26-AUG-1999; 99WO-US019460.  
 XX  
 XX 31-AUG-1998; 98US-0098401P.  
 XX  
 XX (GENTH ) GENENTECH INC.  
 XX  
 XX Henzel WJ;  
 XX WPI; 2000-237934/20.  
 DR  
 XX  
 XX  
 XX Apparatus for protein or polypeptide sample sequencing, uses single  
 PT reagent delivery device to supply chemicals to series of sample  
 PT cartridges.  
 XX  
 XX Example 4; Page 36; 66pp; English.  
 PS  
 CC The invention relates to a novel apparatus for polypeptide sequencing  
 CC which uses a single reagent delivery device to supply chemicals to a  
 CC series of sample cartridges. Each sample cartridge has an inlet and an  
 CC outlet. It is held horizontally in a holder at one of a series of spaced  
 CC holding stations, each including a heater. A valve selectively directs  
 CC chemicals from a single reagent delivery device to sample cartridges. The  
 CC valve also provides connections so that residues from each cartridge can  
 CC be delivered to an analyzer. The apparatus preferably uses a rapid,  
 CC twenty minute, Edman degradation cycle coupled to a rapid and sensitive  
 CC chromatographic separation of amino acid derivatives to maximize sample  
 CC analysis throughput. Rapid and convenient sequencing is achieved using an  
 CC automated and highly accurate device. A single reagent delivery device  
 CC supplies chemicals to a number of sample cartridges. This reduces the  
 CC cost of manufacture and of use. Sequences AAY81463-Y81475 represent the N  
 CC -terminal sequences of a variety of proteins obtained using the apparatus  
 CC of the invention  
 CC  
 XX  
 XX Sequence 10 AA;  
 SQ  
 Query Match 53.8%; Score 28; DB 3; Length 10;  
 Best Local Similarity 71.4%; Pred. No. 1.9e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 LPIEYGP 8  
 DB 1 LPIEYGP 7  
 RESULT 7  
 AAR98051  
 ID AAR98051 standard; peptide; 8 AA.  
 XX  
 XX AAR98051;  
 AC  
 XX  
 XX 27-AUG-2003 (revised)  
 DT 01-AUG-1996 (first entry)  
 XX  
 XX Pseudosteralin G, a cyclic octapeptide.  
 DE  
 XX tyrosinase inhibitor; melanin; skin cosmetic.  
 KM  
 XX Caryophyllaceae sp.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH 1  
 PT Modified-site /note= "not an N-terminal; this amino acid condenses with  
 PT the C-terminal to form a cyclic peptide"  
 FT Modified-site 8

FT /note= "not a C-terminal; see above"  
 XX JP07324095-A.  
 PN 12-DEC-1995.  
 XX  
 XX 30-MAY-1994; 94JP-00116499.  
 PD 30-MAY-1994; 94JP-00116499.  
 XX  
 XX 30-MAY-1994; 94JP-00116499.  
 XX  
 XX (ITOG/) ITOGAMA H.  
 PA (HONS ) YAKULT HONSHA KK.  
 XX  
 XX WPI; 1996-065475/07.  
 DR  
 XX  
 XX New cyclic peptide(s) designated pseudosteralin A-G - are tyrosinase  
 PT inhibitors and melanin formation inhibitors useful in skin cosmetics.  
 PT  
 XX  
 XX Claim 7; Page 2; 13pp; Japanese.  
 PS  
 XX  
 XX The cyclic peptide is one component of the peptides designated  
 CC pseudosteralins A-G which have been extracted from Ratsujin root. These  
 CC peptides are tyrosinase inhibitors and melanin formation inhibitors which  
 CC are useful in skin cosmetics. Pseudosteralins A, B, C, D, E, F and G  
 CC demonstrated tyrosinase inhibiting IC50 values of 131, 187, 63, 100, 175,  
 CC 50 and 75 micron resp., compared with a value of 1.2 mM for arbutin.  
 CC (Updated on 27-AUG-2003 to correct OS field.)  
 CC  
 XX  
 XX Sequence 9 AA;  
 SQ  
 Query Match 51.9%; Score 27; DB 2; Length 8;  
 Best Local Similarity 57.1%; Pred. No. 1.7e+06;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 PLEYGPL 9  
 DB 1 PLEYGPL 7  
 RESULT 8  
 AAB47584  
 ID AAB47584 standard; peptide; 8 AA.  
 XX  
 XX AAB47584;  
 AC  
 XX  
 XX 13-DEC-2001 (first entry)  
 DT  
 XX  
 XX Ag85 complex derived peptide #15.  
 DE  
 XX  
 XX Antigen 85; CD8 T-cell; vaccine; mycobacterium; detection; epitope;  
 KM  
 XX  
 XX Mycobacterium tuberculosis.  
 OS  
 XX  
 XX WO200170991-A1.  
 PN  
 XX  
 XX 27-SEP-2001.  
 PD  
 XX  
 XX 20-MAR-2001; 2001WO-GB001210.  
 PF  
 XX  
 XX 20-MAR-2000; 2000GB-00006693.  
 PR  
 XX  
 XX (GLAX ) GLAXO GROUP LTD.  
 PA  
 XX  
 XX Klein MR;  
 FI  
 XX  
 XX WPI; 2001-611506/70.  
 DR  
 XX  
 XX Use of a polypeptide comprising CD8 T-cell groups derived from  
 PT Mycobacterium tuberculosis antigen 85 in manufacture of a medicament for  
 PT vaccinating prophylactically or therapeutically against mycobacterial  
 PT infection.  
 FT

PS Disclosure; Fig 1; 56pp; English.

CC The sequences given in ABA47570-86 are polypeptides derived from M. tuberculosis antigen 85 which can be recognized by a CD8 T-cell and which may be used in the manufacture of agent for vaccinating prophylactically or therapeutically against infection by a Mycobacterium by stimulating a CD8 T cell response. These polypeptides are useful for vaccinating a pre-selected host to stimulate a CD8 T cell response against a Mycobacterial infection. They are also useful for detecting in a population of T cells the presence or absence of CD8 T cells that recognize an epitope.

CC sequence, by contacting the population of cells comprising CD8 T-cells with the peptide in question and detecting whether the CD8 T-cells recognize the peptide, by detecting the expression of a substance by the T cell, which indicates that the T cell have recognized the peptide or by detecting lysis by T cells of cells that present the peptide on their surface, the detection of lysis indicating that the T cells have recognized the peptide

CC SQ Sequence 8 AA;

Query Match 50.0%; Score 26; DB 4; Length 8;  
Best Local Similarity 80.0%; Pred. No. 1.7e+06;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPITY 6  
|||  
Db 4 LPVEY 8

RESULT 9  
ADE54188  
ID ADE54188 standard; peptide; 8 AA.

AC ADE54188;  
XX  
XX 26-FEB-2004 (first entry)  
XX  
DE Rhesus monkey rotavirus VP8 151-158 amino acid sequence SEQ ID NO:5.  
XX  
XX rotavirus; VP8; rotavirus protein VP4; cytoskeletal; vaccine; cancer;  
KW metastasis inhibition; tumour; inflammation.  
XX  
OS Rhesus rotavirus.  
XX  
XX WO2003098991-A2.  
XX  
XX 04-DEC-2003.  
XX  
XX 10-JAN-2003; 2003WO-IB000280.  
XX  
XX 10-JAN-2003; 2003WO-IB000280.  
XX  
XX (INVE-) CENT INVESTIGACION & ESTUDIOS AVANZADOS.  
XX  
XX Gonzalez-Mariscal L, Nava-Dominguez P;  
XX  
XX WPI; 2004-035006/03.  
XX  
XX Use of rotavirus protein VP4, variants, derivatives or functional  
PT peptides, for preparing a pharmaceutical composition enhancing the  
PT delivery of pharmaceutical agents through the paracellular pathway, or  
PT for treating cytostatic.  
XX  
XX Claim 40; SEQ ID NO 5; 63pp; English.

CC The present invention describes the use of rotavirus protein VP4, its functional variants, derived proteins, derived fusion proteins and functional peptides derived from them as well as their mixtures, for preparing a pharmaceutical composition allowing the passage or enhancing the delivery of pharmaceutical agents through the paracellular pathway.

CC Rotavirus protein VP4 has cytostatic activity, and can be used in vaccine. The rotavirus protein VP4, its functional variants, derived proteins, derived fusion proteins, functional peptides derived from them,

CC or their mixtures is useful for preparing a pharmaceutical composition for the delivery of a therapeutic agent to a subject in need or compositions allowing the passage or enhancing the delivery of pharmaceutical agents through the paracellular pathway. The rotavirus protein is useful for treating cancer, inhibiting metastasis, or for reducing unwanted cellular adhesion occurring between tumour cells or normal cells, as a result of surgery, injury, chemotherapy, disease, inflammation or other pathological condition. Monoclonal or polyclonal antibodies against proteins VP4, VP8 or their derived functional peptides, can be used for exerting effects on the paracellular barrier as proteins VP4, VP8 or their functional derivatives. The rotavirus protein is particularly useful for inducing the opening of tight junctions and increase the paracellular permeability of epithelia and endothelia. The present sequence represents the rhesus monkey rotavirus VP8 151-158 amino acid sequence, which is used in the exemplification of the present invention.

CC SQ Sequence 8 AA;

Query Match 50.0%; Score 26; DB 8; Length 8;  
Best Local Similarity 80.0%; Pred. No. 1.7e+06;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 EYGPL 9  
|||||  
Db 4 QYGPL 8

RESULT 10  
AAV73258  
ID AAV73258 standard; peptide; 9 AA.

AC AAV73258;  
XX  
XX 28-FEB-2000 (first entry)  
XX  
DE HIV-derived MHC class I (CTL) epitope, #416.  
XX  
XX Chimeric; pan DR epitope; expression vector; promoter;  
KW major histocompatibility complex; MHC; targeting; peptide; epitope;  
KW antigen; presentation; class I; cytosolic pathway; endoplasmic reticulum;  
KW class II; extracellular antigen; endocytic pathway; helper T lymphocyte;  
KW HTL; universal epitope; cytotoxic T lymphocyte; CTL; immune response;  
KW immunogenicity; assay; vaccine; immunity; infection; pathogen; virus;  
KW HIV; HBV; HCV; hepatitis B; hepatitis C; bacterium; protozoan;  
KW tumour cell; autoimmune disease; activation; antiviral; antimarialar;  
KW immunoprotective.  
XX  
XX Synthetic.  
OS Human immunodeficiency virus.  
XX  
XX WO9958658-A2.  
XX  
XX 18-NOV-1999.  
XX  
XX 13-MAY-1999; 99WO-US010646.  
XX  
XX 13-MAY-1998; 98US-00078904.  
XX  
XX 15-MAY-1998; 98US-0085751P.  
XX  
XX (EPIM-) EPIMUNE INC.  
XX  
XX Fikes JD, Hermanson GG, Sette A, Ishioka GY, Livingston B;  
PI Chesenut KW;  
XX  
XX WPI; 2000-039103/03.  
XX  
XX Expression vectors encoding major histocompatibility targeting sequence,  
PT used as, e.g. tumor vaccines.  
XX  
XX Claim 11; Page 73; 130pp; English.

CC Sequences AAV73175-Y73266 represent human immunodeficiency virus (HIV) -

CC derived MHC class I (CTL) epitopes which are claimed for use in the  
 CC present invention. The invention relates to a novel expression vector  
 CC comprising a promoter operably linked to a fusion gene encoding a major  
 CC histocompatibility complex (MHC) targeting sequence, and two or more  
 CC heterologous peptide epitopes. The MHC targeting sequence may be a class  
 CC I targeting sequence, which directs an MHC class I epitope to a  
 CC cytosolic pathway or to the endoplasmic reticulum, or an MHC class II  
 CC targeting sequence, which directs extracellular antigens to enter the  
 CC endocytic pathway to be processed into antigen peptides for presentation  
 CC on MHC class II molecules. The heterologous epitopes may comprise either  
 CC helper T lymphocyte (HTL) epitopes, or a cytotoxic T lymphocyte (CTL)  
 CC epitope and a universal HTL epitope such as a pan DR epitope (PADAP). The  
 CC vectors are useful for stimulating an immune response in vivo, as well as  
 CC for use in assaying the human immunogenicity of a human T cell peptide  
 CC epitope in vivo in a non-human mammal. They provide a nucleic acid  
 CC vaccine for enhancing immunity against infectious pathogens, such as  
 CC viruses (e.g., HIV, hepatitis B (HBV) and hepatitis C (HCV)) bacteria,  
 CC protozoa (e.g., Plasmodium falciparum, the cause of malaria) and also  
 CC tumor cells and autoimmune diseases. Universal MHC class II epitopes are  
 CC advantageously combined with other MHC class I and class II epitopes to  
 CC increase the number of cells that are activated in response to a given  
 CC antigen and provide a broader population coverage of MHC-reactive alleles

XX Sequence 9 AA;

Query Match 50.0%; Score 26; DB 3; Length 9;  
 Best Local Similarity 37.5%; Pred. No. 1.7e+06;  
 Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 3 LPVEY 10  
 DB 2 PVHAGPIL 9

RESULT 11  
 AAB82787  
 ID AAB82787 standard; peptide; 9 AA.

XX AAB82787;

DT 29-OCT-2001 (first entry)

XX Epitope of mycobacterial antigen 85A.

XX Mycobacterial antigen 85A; Ag85A; infection; vaccine; diagnosis; therapy.

XX Mycobacterium tuberculosis.

XX WO200158461-A1.

XX 16-AUG-2001.

XX 12-FEB-2001; 2001WO-GB0000561.

XX 10-FEB-2000; 2000GB-00003082.

XX (GLAX ) GLAXO GROUP LTD.

XX Dockrell HM, Smith SM, Brookes R;

XX WPI; 2001-536505/59.

XX Use of polypeptides comprising groups of mycobacterial antigen 85A  
 PT protein and expression vectors comprising polynucleotides encoding the  
 PT polypeptide for vaccination against Mycobacterium infection.

XX Disclosure; Page 5; 49pp; English.

XX The present sequence is that of an epitope peptide comprising amino acid  
 CC residues 48-56 of Mycobacterium tuberculosis antigen 85A (Ag85A).  
 CC Epitopes, including the present sequence, have been found in Ag85A which  
 CC cause the generation of a strong CD8 T-cell response in humans. The CD8 T  
 CC cells which recognise the epitopes are able to lyse macrophages infected

CC with live Mycobacterium tuberculosis. A polypeptide which comprises an  
 CC epitope structure of Ag85A, or an expression vector comprising a  
 CC polynucleotide encoding such a polypeptide, is used in the manufacture  
 CC of a medicament for vaccinating prophylactically or therapeutically against  
 CC infection by a mycobacterium stimulating a CD8 T-cell response. The  
 CC polypeptide or expression vector may be within an antigen-presenting  
 CC cell. A claimed vaccine composition comprises the polypeptide or  
 CC expression vector and an adjuvant or delivery system capable of  
 CC stimulating a CD8 T-cell response. Also claimed are methods of detecting  
 CC in a population of T-cells the presence or absence of CD8 T-cells that  
 CC recognise the epitope sequence, a method of diagnosing mycobacterial  
 CC infection or of testing the effectiveness of a vaccination, and a method  
 CC of treating a mycobacterium infection by administering T-cells capable of  
 CC recognising the epitope

XX Sequence 9 AA;

Query Match 50.0%; Score 26; DB 4; Length 9;  
 Best Local Similarity 80.0%; Pred. No. 1.7e+06;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 LPVEY 6  
 DB 2 LPVEY 6

RESULT 12  
 AAB47585  
 ID AAB47585 standard; peptide; 9 AA.

XX AAB47585;

DT 13-DEC-2001 (first entry)

XX Ag85 complex derived peptide #16.

XX Antigen 85; CD8 T-cell; vaccine; mycobacterium; detection; epitope;

XX lysis.

XX Mycobacterium tuberculosis.

XX WO200170991-A1.

XX 27-SEP-2001.

XX 20-MAR-2001; 2001WO-GB001210.

XX 20-MAR-2000; 2000GB-00006693.

XX (GLAX ) GLAXO GROUP LTD.

XX Klein MR;

XX WPI; 2001-611506/70.

XX Use of a polypeptide comprising CD8 T-cell groups derived from  
 PT Mycobacterium tuberculosis antigen 85 in manufacture of a medicament for  
 PT vaccinating prophylactically or therapeutically against mycobacterial  
 PT infection.

XX Disclosure; Fig 1; 56pp; English.

XX The sequences given in AAB47570-86 are polypeptides derived from M.  
 CC tuberculosis antigen 85 which can be recognized by a CD8 T-cell and which  
 CC may be used in the manufacture of agent for vaccinating prophylactically  
 CC or therapeutically against infection by a mycobacterium by stimulating a  
 CC CD8 T cell response. These polypeptides are useful for vaccinating a pre-  
 CC selected host to stimulate a CD8 T cell response against a Mycobacterial  
 CC infection. They are also useful for detecting in a population of T cells  
 CC the presence or absence of CD8 T cells that recognize an epitope  
 CC sequence, by contacting the population of cells comprising CD8 T-cells  
 CC with the peptide in question and detecting whether the CD8 T-cells  
 CC recognize the peptide, by detecting the expression of a substance by the

CC T cell which indicates that the T cell have recognized the peptide or by  
 CC detecting lysis by T cells of cells that present the peptide on their  
 CC surface, the detection of lysis indicating that the T cells have  
 CC recognized the peptide

XX Sequence 9 AA;

Query Match 50.0%; Score 26; DB 4; Length 9;  
 Best Local Similarity 80.0%; Pred. No. 1.7e+06;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPEY 6  
 DB 4 LPEY 8

RESULT 13  
 AAU98288 standard; peptide; 9 AA.

XX AAU98288;

XX 15-AUG-2002 (first entry)

XX M. tuberculosis Ag85 derived immunodominant T cell epitope #15.  
 XX Ag85; epitope; vaccine; tuberculosis; mycolyltransferase; HLA-A\*02001;  
 XX Mycobacterium tuberculosis; Mycobacterium leprae; Mycobacterium bovis;  
 XX Mycobacterium ulcerans; Mycobacterium avium; anti-bacterial;  
 XX MHC class I-restricted CD8+ T cell; human leukocyte antigen.

XX Mycobacterium tuberculosis.

XX Key Location/Qualifiers

XX Region 2 /note="Amino acid contributing to the HLA-A\*02001  
 binding motif"  
 XX Region 9 /note="Amino acid contributing to the HLA-A\*02001  
 binding motif"

XX EPI211260-A1.

XX 05-JUN-2002.

XX 30-NOV-2000; 2000EP-00204268.

XX 30-NOV-2000; 2000EP-00204268.

XX (ZIEK-) ACAD ZIEKENHUIS LEIDEN.

XX Ottenhof THM, Geluk A;

XX WPI; 2002-464923/50.

XX Peptide of 8-11 amino acids derived from the Ag85 protein of  
 PT mycobacterium induce proliferation of MHC class I-restricted CD8+ T cells  
 PT and are useful to vaccinate against infection by mycobacterium.

XX Claim 1; Page 7; 20pp; English.

XX The invention relates to peptide (P1) derived from an Ag85 protein of  
 CC mycobacterium (which is associated with mycolyltransferase activity and  
 CC is involved in cell wall synthesis), comprising 8-11 amino acids, and  
 CC capable of inducing proliferation of MHC class I-restricted CD8+ T cells  
 CC in vitro. Also included are (1) a peptide comprising P1 flanked by amino  
 CC acids representing antigen processing sites; (2) a polypeptide comprising  
 CC at least two of the peptides in (1); (3) a nucleic acid encoding one of  
 CC the claimed peptides or polypeptides; (4) a vector comprising the above  
 CC nucleic acid; (5) a host cell comprising the above nucleic acid or vector  
 CC; and (6) detecting and/or enumerating CD8+ T cells against  
 CC mycobacterium, comprising tetrameric complexes of MHC (major  
 CC histocompatibility group) class I and one of the claimed peptides or

CC polypeptides. The molecules of the invention are used to prepare a  
 CC vaccine against mycobacterium (e.g. Mycobacterium tuberculosis,  
 CC Mycobacterium leprae, Mycobacterium bovis, Mycobacterium ulcerans,  
 CC Mycobacterium avium) and thus protect against diseases such as  
 CC tuberculosis. The present sequence is an M. tuberculosis Ag85 derived  
 CC immunogenic peptide of the invention binding to HLA-A\*02001 (human  
 CC leukocyte antigen)

XX Sequence 9 AA;

Query Match 50.0%; Score 26; DB 5; Length 9;  
 Best Local Similarity 80.0%; Pred. No. 1.7e+06;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPEY 6  
 DB 2 LPEY 6

RESULT 14  
 AB000162 standard; peptide; 9 AA.

XX AB000162;

XX 29-AUG-2003 (revised)

XX 30-AUG-2002 (first entry)

XX HIV GAG protein derived HLA antigen.

XX Human major histocompatibility complex; HLA; immunogen; cancer;  
 XX viral infection; vaccine; antigen; anti-HIV; virucide; cytostatic;  
 XX protozoicide.

XX Human immunodeficiency virus 1.

XX WO200220035-A1.

XX 14-MAR-2002.

XX 01-SEP-2000; 2000WO-US023913.

XX 01-SEP-2000; 2000WO-US023913.

XX (EPIM-) EPIMUNE INC.

XX Sette A, Sidney J, Southwood S;

XX WPI; 2002-479452/51.

XX New immunogenic peptide with B7-like supermotif, useful in vaccines  
 PT against e.g. viral infection and cancer, induces a cytotoxic T cell  
 PT response.

XX Example 1; Page 29; 39pp; English.

XX The present invention relates to compositions containing immunogenic  
 CC peptides, which comprise a B7-like supermotif. The compositions can be  
 CC used, as vaccines, to treat or prevent viral diseases (hepatitis B or C,  
 CC Epstein-Barr, human immune deficiency, Lassa fever or cytomegalovirus),  
 CC cancers (e.g. of prostate, kidney or cervix, or lymphoma, where  
 CC associated with expression of p53, carcino-embryonal antigen or  
 CC Her2/neu), or malaria. The peptides are also useful as diagnostic agents,  
 CC e.g. to predict the outcome of a particular therapy and to identify  
 CC subjects at risk of developing a chronic infection. Nucleic acids  
 CC encoding the peptides can be used in DNA vaccines. The present sequence  
 CC is an immunogenic peptide of the invention. (Updated on 29-AUG-2003 to  
 CC standardise OS field)

XX Sequence 9 AA;

Query Match 50.0%; Score 26; DB 5; Length 9;  
 Best Local Similarity 37.5%; Pred. No. 1.7e+06;

Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 PIRXGPLV 10  
 | : | : | :  
 Db 2 PVHAGPIL 9

Db 2 PVHAGPIL 9  
 Search completed: November 30, 2004, 08:09:37  
 Job time : 57.333 secs

## RESULT 15

ABO01027  
 ID ABO01027 standard; peptide; 9 AA.  
 XX  
 AC ABO01027;  
 XX  
 DT 23-OCT-2003 (revised)  
 DT 05-AUG-2003 (first entry)  
 XX  
 DE B7-1-like supermotif-containing, immunogenic peptide 77.  
 XX  
 KM B7-1-like peptide; immunogen; HLA; human leukocyte antigen; virucide;  
 KM hepatocytic; antiinflammatory; anti-HIV; cytostatic; immunostimulant;  
 KM viral infection; cancer; hepatitis B virus infection; AIDS;  
 KM hepatitis C virus infection; cytomegalovirus infection; prostate cancer;  
 KM acquired immunodeficiency syndrome; renal carcinoma; cervical carcinoma;  
 KM lymphoma; condyloma acuminatum; chronic infection.  
 XX  
 OS Human immunodeficiency virus 1.  
 XX  
 PN US2002177694-A1.  
 XX  
 PD 28-NOV-2002.  
 XX  
 PF 03-FEB-1998; 98US-00017743.  
 XX  
 PR 23-JAN-1996; 96US-00590298.  
 XX  
 PA (SETT/) SETT A.  
 PA (SIDN/) SIDNEY J.  
 PA (SOUT/) SOUTWOOD S.  
 XX  
 PI Sette A, Sidney J, Southwood S;  
 XX  
 DR WPI: 2003-447321/42.  
 XX  
 PT Composition for treating and/or preventing viral infection and cancer  
 PT such as prostate cancer, comprises an immunogenic peptide having a B7-  
 PT like supermotif.  
 XX  
 PS Claim 1; Page 11; 14pp; English.  
 XX  
 CC The invention relates to a composition comprising an immunogenic peptide  
 CC (IP) having a B7-1-like supermotif, where IP is selected from one of the  
 CC 127 peptide sequences (appearing as ABO00951-ABO01082). The composition  
 CC is useful for treating and/or preventing viral infection and cancer.  
 CC including infection by hepatitis B, hepatitis C or cytomegalovirus (CMV),  
 CC prostate cancer, acquired immunodeficiency syndrome (AIDS), renal  
 CC carcinoma, cervical carcinoma, lymphoma and condyloma acuminatum. The  
 CC composition is also useful for eliciting an immune response against a  
 CC desired antigen, in therapeutic or diagnostic applications, for  
 CC preventing the evolution from acute to chronic infection, for the  
 CC treatment of chronic infection, and to stimulate the immune system to  
 CC eliminate virus-infected cells in carriers. The peptides are HLA (human  
 CC leukocyte antigen) binding peptides. The present sequence is an  
 CC immunogenic peptide of the invention, having a B7-1-like supermotif.  
 CC (Updated on 23-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 9 AA;

Query Match 50.0%; Score 26; DB 6; Length 9;  
 Best Local Similarity 37.5%; Pred. No. 1.7e+06;  
 Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 PIRXGPLV 10  
 | : | : | :  
 Db 2 PVHAGPIL 9



APPLICANT: Peng, Yue  
APPLICANT: Tularex Inc.  
TITLE OF INVENTION: Diagnosis and Treatment of Cancer Using Mammalian  
TITLE OF INVENTION: Pellino polypeptides and Polynucleotides  
FILE REFERENCE: 018781-006810US  
CURRENT APPLICATION NUMBER: US/10/041,030  
CURRENT FILING DATE: 2001-12-28  
PRIOR APPLICATION NUMBER: US 60/259,502  
PRIOR FILING DATE: 2001-01-02  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO: 14  
LENGTH: 7  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: amino acid  
OTHER INFORMATION: sequence identity in comparison of pellino 1 and  
US-10-041-030-14

Query Match  
Best Local Similarity 51.9%; Score 27; DB 13; Length 7;  
Best Local Similarity 57.1%; Pred. No. 1,4e+06;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 PIRGPLY 9  
DB 1 PVRIGEL 7

RESULT 3  
US-09-017-743C-77  
Sequence 77, Application US/09017743C  
Patent No. US2002017694A1  
GENERAL INFORMATION:  
APPLICANT: Sette, Alessandro  
Sidney, John  
Southwood, Scott  
TITLE OF INVENTION: HLA Binding Peptides and Their  
Uses  
NUMBER OF SEQUENCES: 146  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/017,743C  
FILING DATE: 03-Feb-1998  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/590,298  
FILING DATE: 23-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Parent, Annette S.  
REGISTRATION NUMBER: 42,058  
REFERENCE/DOCKET NUMBER: 018623-008050US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0200  
TELEFAX: 415-576-0300  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 77:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 77:  
US-09-017-743C-77

Query Match  
Best Local Similarity 50.0%; Score 26; DB 9; Length 9;  
Best Local Similarity 37.5%; Pred. No. 1,4e+06;  
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 PIRGPLY 10  
DB 2 PVRHAGPII 9

RESULT 4  
US-10-371-525-404  
Sequence 404, Application US/10371525  
Publication No. US20030203869A1  
GENERAL INFORMATION:  
APPLICANT: Fikes, John D.  
APPLICANT: Hermanson, Gary G.  
APPLICANT: Sette, Alessandro  
APPLICANT: Ishioka, Glenn Y.  
APPLICANT: Livingston, Brian  
APPLICANT: Chesnut, Robert W.  
APPLICANT: EpiImmune Inc.  
TITLE OF INVENTION: Expression Vectors for Stimulating an  
Immune Response and Methods of Using the Same  
FILE REFERENCE: 39963-20022.01  
CURRENT APPLICATION NUMBER: US/10/371,525  
PRIOR FILING DATE: 2003-02-21  
PRIOR APPLICATION NUMBER: US 09/311,784  
PRIOR FILING DATE: 1999-05-13  
PRIOR APPLICATION NUMBER: US 60/085,751  
PRIOR FILING DATE: 1998-05-15  
NUMBER OF SEQ ID NOS: 463  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 404  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HIV GAG 248 (peptide 1292.14)  
US-10-371-525-404

Query Match  
Best Local Similarity 50.0%; Score 26; DB 14; Length 9;  
Best Local Similarity 37.5%; Pred. No. 1,4e+06;  
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 PIRGPLY 10  
DB 2 PVRHAGPII 9

RESULT 5  
US-10-371-069-404  
Sequence 404, Application US/10371069  
Publication No. US20030216342A1  
GENERAL INFORMATION:  
APPLICANT: EpiImmune Inc.  
APPLICANT: Fikes, John D.  
APPLICANT: Hermanson, Gary G.  
APPLICANT: Sette, Alessandro  
APPLICANT: Ishioka, Glenn Y.  
APPLICANT: Livingston, Brian  
APPLICANT: Chesnut, Robert W.  
APPLICANT: EpiImmune Inc.  
TITLE OF INVENTION: Expression Vectors for Stimulating an  
Immune Response and Methods of Using the Same  
FILE REFERENCE: 39963-20022.10  
CURRENT APPLICATION NUMBER: US/10/371,069  
PRIOR FILING DATE: 2003-02-21  
PRIOR APPLICATION NUMBER: US 09/078,904

PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: US 60/085,751  
PRIOR FILING DATE: 1998-05-15  
NUMBER OF SEQ ID NOS: 463  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO: 404  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HIV GAG 248 (peptide 1292.14)  
US-10-371-069-404

Query Match 50.0%; Score 26; DB 14; Length 9;  
Best Local Similarity 37.5%; Pred. No. 1.4e+06;  
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 PIRYGLV 10  
DB 2 PVHAGPII 9

RESULT 6  
US-10-371-645-404  
Sequence 404, Application US/10371645  
Publication No. US20030216343A1  
GENERAL INFORMATION:  
APPLICANT: EPIMUNE INC.  
APPLICANT: Fikes, John D.  
APPLICANT: Hermanson, Gary G.  
APPLICANT: Settle, Alessandro  
APPLICANT: Ishioaka, Glenn Y.  
APPLICANT: Livingston, Brian  
APPLICANT: Chesnut, Robert W.  
APPLICANT: Epimmune Inc.  
TITLE OF INVENTION: Expression Vectors for Stimulating an  
TITLE OF INVENTION: Immune Response and Methods of Using the Same  
FILE REFERENCE: 39963-20022.11  
CURRENT FILING DATE: 2003-06-20  
PRIOR APPLICATION NUMBER: US/10/371,645  
PRIOR FILING DATE: 2003-06-20  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: US 09/078,904  
PRIOR FILING DATE: 1998-05-15  
NUMBER OF SEQ ID NOS: 463  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO: 404  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HIV GAG 248 (peptide 1292.14)  
US-10-371-645-404

Query Match 50.0%; Score 26; DB 14; Length 9;  
Best Local Similarity 37.5%; Pred. No. 1.4e+06;  
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 PIRYGLV 10  
DB 2 PVHAGPII 9

RESULT 7  
US-10-371-260-404  
Sequence 404, Application US/10371260  
Publication No. US20030220285A1  
GENERAL INFORMATION:  
APPLICANT: EPIMUNE INC.  
APPLICANT: Fikes, John D.  
APPLICANT: Hermanson, Gary G.  
APPLICANT: Settle, Alessandro  
APPLICANT: Ishioaka, Glenn Y.

APPLICANT: Livingston, Brian  
APPLICANT: Chesnut, Robert W.  
APPLICANT: Epimmune Inc.  
TITLE OF INVENTION: Expression Vectors for Stimulating an  
TITLE OF INVENTION: Immune Response and Methods of Using the Same  
FILE REFERENCE: 39963-20022.11  
CURRENT FILING DATE: 2003-02-21  
PRIOR APPLICATION NUMBER: US/10/371,260  
PRIOR FILING DATE: 2003-02-21  
PRIOR APPLICATION NUMBER: US 09/078,904  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: US 60/085,751  
PRIOR FILING DATE: 1998-05-15  
NUMBER OF SEQ ID NOS: 463  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO: 404  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HIV GAG 248 (peptide 1292.14)  
US-10-371-260-404

Query Match 50.0%; Score 26; DB 14; Length 9;  
Best Local Similarity 37.5%; Pred. No. 1.4e+06;  
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 PIRYGLV 10  
DB 2 PVHAGPII 9

RESULT 8  
US-10-203-562-1  
Sequence 1, Application US/10203562  
Publication No. US20040037837A1  
GENERAL INFORMATION:  
APPLICANT: HAZEL M. DOCKRELL  
APPLICANT: Steven M. SMITH  
APPLICANT: Roger BROOKES  
TITLE OF INVENTION: MYCOBACTERIUM AG85 COMPLEX-SPECIFIC T CELL PEPTIDES AND USE IN  
TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC APPLICATIONS THEREOF  
FILE REFERENCE: 117-409 / N.784688 GCW  
CURRENT FILING DATE: 2002-08-12  
PRIOR APPLICATION NUMBER: US/10/203,562  
PRIOR FILING DATE: 2002-08-12  
PRIOR APPLICATION NUMBER: PCT/GB01/00561  
PRIOR FILING DATE: 2001-02-12  
PRIOR APPLICATION NUMBER: GB 0003082.5  
PRIOR FILING DATE: 2000-02-10  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 1  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Mycobacterium tuberculosis  
US-10-203-562-1

Query Match 50.0%; Score 26; DB 15; Length 9;  
Best Local Similarity 80.0%; Pred. No. 1.4e+06;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPREY 6  
DB 2 LPREY 6

RESULT 9  
US-09-572-404B-461  
Sequence 461, Application US/09572404B  
Publication No. US20030078374A1  
GENERAL INFORMATION:  
APPLICANT: Proteom Ltd  
TITLE OF INVENTION: Complementary peptide ligands from the human genome  
FILE REFERENCE: Human patent

CURRENT APPLICATION NUMBER: US/09/572,404B  
CURRENT FILING DATE: 2000-05-17  
NUMBER OF SEQ ID NOS: 4203  
SOFTWARE: ProPatent version 1.0  
SEQ ID NO 461  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Homo Sapiens  
FEATURE:  
OTHER INFORMATION: Sequence located in CLASS2 OR AP17 at 16-25 and may interact with  
US-09-572-404B-461

Query Match  
Best Local Similarity 48.1%; Score 25; DB 10; Length 10;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 LPIEYGP 8  
DB 4 LPSESGP 10

RESULT 10  
US-09-932-165-801  
Sequence 801, Application US/09932165  
Publication No. US20030134784A1  
GENERAL INFORMATION:  
APPLICANT: RAITANO, ARTHUR  
APPLICANT: CHALLITH-ELD, PTA M.  
APPLICANT: FARIS, MARI  
APPLICANT: SAFRAN, DOUGLAS  
APPLICANT: AFAR, DANIEL  
APPLICANT: LEVIN, ELANA  
APPLICANT: HUBERT, RENE  
APPLICANT: GE, WANGMAO  
APPLICANT: JAKBOVITS, AVA  
TITLE OF INVENTION: NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED  
TITLE OF INVENTION: 83P23 AND CATREPELLI USEFUL IN TREATMENT AND  
TITLE OF INVENTION: DETECTION OF CANCER  
FILE REFERENCE: 51158-20014.00  
CURRENT APPLICATION NUMBER: US/09/932,165  
CURRENT FILING DATE: 2001-08-17  
PRIOR APPLICATION NUMBER: 60/226,329  
PRIOR FILING DATE: 2000-08-17  
NUMBER OF SEQ ID NOS: 1508  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 801  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Peptide motif  
US-09-932-165-801

Query Match  
Best Local Similarity 46.2%; Score 24; DB 10; Length 9;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YGPL 9  
DB 2 YGPL 5

RESULT 11  
US-08-344-824-11  
Sequence 11, Application US/08344824  
Publication No. US20030152560A1  
GENERAL INFORMATION:  
APPLICANT: SETTE, Alessandro  
APPLICANT: SIDNEY, John  
TITLE OF INVENTION: H2A BINDING PEPTIDES AND THEIR USES  
NUMBER OF SEQUENCES: 399  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend Hourie and Crew  
STREET: One Market Plaza, Stewart Street Tower, 20th  
STREET: Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/344,824  
FILING DATE: 23-NOV-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/278,634  
FILING DATE: 21-JUL-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 14137-80-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-344-824-11

Query Match  
Best Local Similarity 46.2%; Score 24; DB 8; Length 10;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LPIEYGPV 10  
DB 2 LPSPYPSV 10

RESULT 12  
US-09-898-860-126  
Sequence 126, Application US/09898860  
Publication No. US20030144482A1  
GENERAL INFORMATION:  
APPLICANT: KAKAKI, YUTAKA, ROSENBERG,  
STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
METHODS  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/898,860  
FILING DATE: 03-JUL-2001  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/267,439  
FILING DATE: <Unknown>

APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 126:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 126:  
US-09-898-860-126

Query Match 46.2%; Score 24; DB 10; Length 10;  
Best Local Similarity 55.6%; Pred. No. 1.1e+03;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LPIEYGPLV 10  
DB 2 LPSDYFSPV 10

RESULT 13  
US-09-935-476-4  
Sequence 4, Application US/09935476  
Publication No. US20040096445A1  
GENERAL INFORMATION:  
APPLICANT: Epimmune, Inc.  
APPLICANT: Sidney, John  
APPLICANT: Sette, Alessandro  
APPLICANT: Grey, Howard  
APPLICANT: Southwood, Scott  
TITLE OF INVENTION: SUBUNIT VACCINES WITH A2 SUPERMOTIFS  
FILE REFERENCE: 39963-20029.20  
CURRENT APPLICATION NUMBER: US/09/935,476  
CURRENT FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: US 09/346,105  
PRIOR FILING DATE: 1999-06-30  
PRIOR APPLICATION NUMBER: US 60/264,969  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Analog of the HBV core 18027 epitope  
US-09-935-476-4  
Query Match 46.2%; Score 24; DB 11; Length 10;  
Best Local Similarity 55.6%; Pred. No. 1.1e+03;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LPIEYGPLV 10  
DB 2 LPSDYFSPV 10

RESULT 14  
US-10-094-699-108  
Sequence 108, Application US/10094699  
Publication No. US20030046714A1  
GENERAL INFORMATION:

APPLICANT: SIMARD, John, J.L.  
APPLICANT: DIAMOND, David, C.  
TITLE OF INVENTION: ANTI-NEOVASCULATURE PREPARATIONS FOR  
FILE REFERENCE: CTLMW.015A  
CURRENT APPLICATION NUMBER: US/10/094,699  
CURRENT FILING DATE: 2002-06-21  
PRIOR APPLICATION NUMBER: US 60/274,063  
PRIOR FILING DATE: 2001-03-07  
NUMBER OF SEQ ID NOS: 109  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 108  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Homo sapien  
US-10-094-699-108  
Query Match 46.2%; Score 24; DB 14; Length 10;  
Best Local Similarity 55.6%; Pred. No. 1.1e+03;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LPIEYGPLV 10  
DB 2 LPSDYFSPV 10

RESULT 15  
US-10-359-431-25  
Sequence 25, Application US/10359431  
Publication No. US20030171538A1  
GENERAL INFORMATION:  
APPLICANT: Chiesari, Francis V.  
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T  
Lymphocyte Responses to Hepatitis B Virus  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/359,431  
FILING DATE: 05-Feb-2003  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,540  
FILING DATE: 26-AUG-1991  
APPLICATION NUMBER: US 07/935,898  
FILING DATE: 26-AUG-1992  
APPLICATION NUMBER: US 08/100,870  
FILING DATE: 02-AUG-1993  
APPLICATION NUMBER: WO PCT/US94/08685  
FILING DATE: 01-AUG-1994  
APPLICATION NUMBER: US 08/591,502  
FILING DATE: 20-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Lockyer, Jean M.  
REGISTRATION NUMBER: 14,879  
REFERENCE/DOCKET NUMBER: 014740-000231US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid

Tue Nov 30 08:49:25 2004

us-10-008-377a-6.closed.rapb

Page 6

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 25:  
US-10-359-431-25

Query Match 46.2% Score 24; DB 14; Length 10;  
Best Local Similarity 55.6%; Pred. No. 1.1e+03;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 LPIEYGPLV 10  
Db 2 LPSDYFESV 10

Search completed: November 30, 2004, 08:31:01  
Job time : 45 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:41:57 ; Search time 27 Seconds  
(without alignments)

24.562 Million cell updates/sec

Title: US-10-008-377a-6

Perfect score: 52

Sequence: 1 ADPIEXGPV 10

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 110780

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: /cgn2\_6/prodata/1/aa/5A.COMB.pep.\*  
2: /cgn2\_6/prodata/1/aa/5B.COMB.pep.\*  
3: /cgn2\_6/prodata/1/aa/6A.COMB.pep.\*  
4: /cgn2\_6/prodata/1/aa/6B.COMB.pep.\*  
5: /cgn2\_6/prodata/1/aa/PCUS.COMB.pep.\*  
6: /cgn2\_6/prodata/1/aa/backfilest.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	65.4	10	3	US-09-530-058-3
2	27	51.9	7	3	US-08-786-45B-3
3	26	50.0	9	4	US-09-311-784A-404
4	26	50.0	10	2	US-08-733-505A-32
5	26	50.0	10	2	US-08-706-741B-67
6	26	50.0	10	2	US-08-924-655A-67
7	26	50.0	10	3	US-09-001-984C-92
8	26	50.0	10	3	US-09-001-984C-97
9	26	50.0	10	3	US-09-001-984C-101
10	26	50.0	10	4	US-09-396-347F-92
11	26	50.0	10	4	US-09-396-347F-97
12	26	50.0	10	4	US-09-396-347F-101
13	25	48.1	7	3	US-08-604-965E-10
14	25	48.1	7	3	US-08-604-965E-15
15	25	48.1	7	6	5258287-43
16	25	48.1	8	6	5258287-44
17	25	48.1	10	4	US-09-239-043D-2314
18	24	46.2	9	1	US-08-276-452A-58
19	24	46.2	9	1	US-08-798-744-58
20	24	46.2	10	1	US-08-195-075-12
21	24	46.2	10	1	US-08-363-691-9
22	24	46.2	10	2	US-08-417-174-126
23	24	46.2	10	3	US-08-155-359A-647
24	24	46.2	10	3	US-09-267-439-126
25	24	46.2	10	3	US-08-992-877-1
26	24	46.2	10	4	US-09-311-784A-226
27	24	46.2	10	4	US-09-073-138-126

28	24	46.2	10	4	US-09-497-233-4	Sequence 4, Appli
29	24	46.2	10	4	US-09-543-608A-43	Sequence 43, Appl
30	24	46.2	10	4	US-08-591-502B-25	Sequence 25, Appl
31	24	46.2	10	4	US-08-234-784B-97	Sequence 97, Appl
32	24	46.2	10	4	US-09-239-043D-2254	Sequence 2254, Ap
33	24	46.2	10	4	US-09-239-043D-2307	Sequence 2307, Ap
34	24	46.2	10	4	US-09-239-043D-2313	Sequence 2313, Ap
35	24	46.2	10	4	US-09-239-043D-2313	Sequence 2313, Ap
36	24	46.2	10	5	PCT-US95-01672-12	Sequence 12, Appl
37	23	44.2	8	3	US-08-444-818-417	Sequence 417, Appl
38	23	44.2	9	1	US-08-615-181-18	Sequence 18, Appl
39	23	44.2	9	4	US-09-311-784A-370	Sequence 370, Appl
40	23	44.2	9	4	US-09-820-174-33	Sequence 33, Appl
41	23	44.2	9	4	US-09-820-195A-33	Sequence 33, Appl
42	23	44.2	10	3	US-09-561-366B-33	Sequence 33, Appl
43	23	44.2	10	4	US-10-114-176-33	Sequence 33, Appl
44	23	44.2	10	4	US-09-239-043D-2288	Sequence 2288, Ap
45	23	44.2	10	4	US-09-239-043D-2308	Sequence 2308, Ap

#### ALIGNMENTS

RESULT 1  
US-09-530-058-3  
; Sequence 3, Application US/09530058  
; Patent No. 637938  
GENERAL INFORMATION:

APPLICANT: <Unknown>

TITLE OF INVENTION: Epoxide hydrolase

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Knobbe, Martens, Olson & Bear

STREET: 501 West Broadway

CITY: San Diego

STATE: California

COUNTRY: USA

ZIP: 92101-3505

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: Win95

SOFTWARE: Word

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/530,058

FILING DATE: 03-Jul-2000

ATTORNEY/AGENT INFORMATION:

NAME: Harte, Daniel

REGISTRATION NUMBER: 40,637

REFERENCE/DOCKET NUMBER: VANM150.001APC

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 10 amino acids

TYPE: amino acid

STRANDEDNESS: single

MOLECULE TYPE: linear

HYPOTHEICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-530-058-3

Query Match 65.4%; Score 34; DB 3; Length 10;

Best Local Similarity 63.3%; Pred. No. 5.2;

Matches 5; Conservative 1; Mismatches 0; Gaps 0;

QV 3 PILEUP 8

DB 2 PILEUP 7

RESULT 2  
US-08-786-455B-3  
Sequence 3, Application US/08786455B  
Patent No. 6193971  
GENERAL INFORMATION:  
APPLICANT: HOFMANN, Joachim  
APPLICANT: SCHMID, Karlheinz  
TITLE OF INVENTION: DICTYOCALUS VITIPARUS ANTIGEN FOR  
TITLE OF INVENTION: DIAGNOSING LUNGWORM INFESTATION AND FOR VACCINATION  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESS: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/786,455B  
FILING DATE: 21-JAN-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 196 01 754.8  
FILING DATE: 19-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: GRANADOS, Patricia D.  
REGISTRATION NUMBER: 33,683  
REFERENCE/DOCKET NUMBER: 18748/327  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 672-5300  
TELEFAX: (202) 672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-786-455B-3  
Query Match  
Best Local Similarity 51.9%; Score 27; DB 3; Length 7;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 PLEY 6  
DB 1 PLEY 5  
RESULT 3  
US-09-311-784A-404  
Sequence 404, Application US/09311784A  
Patent No. 6534482  
GENERAL INFORMATION:  
APPLICANT: Fikes, John D.  
APPLICANT: Hermannson, Gary G.  
APPLICANT: Sette, Alessandro  
APPLICANT: Ishioka, Glenn Y.  
APPLICANT: Livingston, Brian  
APPLICANT: Chesnut, Robert W.  
APPLICANT: Epiimmune, Inc.  
TITLE OF INVENTION: Expression Vectors for Stimulating an  
TITLE OF INVENTION: Immune Response and Methods of Using the Same  
FILE REFERENCE: 39963-20022.01  
CURRENT APPLICATION NUMBER: US/09/311,784A  
CURRENT FILING DATE: 1999-05-13

PRIOR APPLICATION NUMBER: US 60/085,751  
PRIOR FILING DATE: 1998-05-15  
NUMBER OF SEQ ID NOS: 463  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 404  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HIV GAG 248 (peptide 1292.14)  
US-09-311-784A-404  
Query Match  
Best Local Similarity 50.0%; Score 26; DB 4; Length 9;  
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
QY 3 PLEY 10  
DB 2 PMSYGR 9  
RESULT 4  
US-08-733-505A-32  
Sequence 32, Application US/08733505A  
Patent No. 5856445  
GENERAL INFORMATION:  
APPLICANT: KORSMEYER, STANLEY J.  
TITLE OF INVENTION: SERINE SUBSTITUTED MUTANTS OF  
TITLE OF INVENTION: BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: HOWELL & HAFFERKAMP, L.C.  
STREET: 7733 FORSYTH BLVD., SUITE 1400  
CITY: ST. LOUIS  
STATE: MISSOURI  
COUNTRY: USA  
ZIP: 63105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/733,505A  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: HOLLAND, DONALD R.  
REGISTRATION NUMBER: 35,197  
REFERENCE/DOCKET NUMBER: 965458  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 727-5188  
TELEFAX: (314) 727-6092  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-733-505A-32  
Query Match  
Best Local Similarity 50.0%; Score 26; DB 2; Length 10;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 3 PLEY 10  
DB 2 PMSYGR 9  
RESULT 5  
US-08-706-741B-67

Sequence 67, Application US/08706741B  
Patent No. 5955593  
GENERAL INFORMATION:  
APPLICANT: KORSMEYER, STANLEY J.  
TITLE OF INVENTION: BH3 INTERACTING DOMAIN DEATH AGONIST  
NUMBER OF SEQUENCES: 88  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
STREET: 7733 FORSYTH BLVD., SUITE 1400  
CITY: ST. LOUIS  
STATE: MISSOURI  
COUNTRY: USA  
ZIP: 63146  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/706,741B  
FILING DATE: 09-SEP-1996  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: HOLLAND, DONALD R.  
REGISTRATION NUMBER: 35,197  
REFERENCE/DOCKET NUMBER: 965017  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 727-5188  
TELEFAX: (314) 727-6092  
INFORMATION FOR SEQ ID NO: 67:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-706-741B-67

Query Match  
Best Local Similarity 50.0%; Score 26; DB 2; Length 10;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 PLEYGPLV 10  
DB 2 FMSYGRLL 9

RESULT 6  
US-08-924-695A-67  
Sequence 67, Application US/08924695A  
Patent No. 5985583  
GENERAL INFORMATION:  
APPLICANT: KORSMEYER, STANLEY J.  
TITLE OF INVENTION: BH3 INTERACTING DOMAIN DEATH AGONIST  
NUMBER OF SEQUENCES: 88  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
STREET: 7733 FORSYTH BLVD., SUITE 1400  
CITY: ST. LOUIS  
STATE: MISSOURI  
COUNTRY: USA  
ZIP: 63105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/924,695A  
FILING DATE: 09-SEP-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: HOLLAND, DONALD R.

REGISTRATION NUMBER: 35,197  
REFERENCE/DOCKET NUMBER: 971798  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 727-5188  
TELEFAX: (314) 727-6092  
INFORMATION FOR SEQ ID NO: 67:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-924-695A-67

Query Match  
Best Local Similarity 50.0%; Score 26; DB 2; Length 10;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 PLEYGPLV 10  
DB 2 FMSYGRLL 9

RESULT 7  
US-09-001-984C-92  
Sequence 92, Application US/09001984C  
Patent No. 6245331  
GENERAL INFORMATION:  
APPLICANT: Laal, Susan  
APPLICANT: Zolla-Pazner, Susan  
TITLE OF INVENTION: EARLY DETECTION OF MYCOBACTERIAL DISEASE  
FILE REFERENCE: NYU-011  
CURRENT APPLICATION NUMBER: US/09/001,984C  
CURRENT FILING DATE: 1997-12-31  
PRIOR APPLICATION NUMBER: 60/034,003  
PRIOR FILING DATE: 1996-12-31  
NUMBER OF SEQ ID NOS: 106  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 92  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Mycobacterium tuberculosis strain H37Rv  
US-09-001-984C-92

Query Match  
Best Local Similarity 50.0%; Score 26; DB 3; Length 10;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPYEV 6  
DB 6 LPYEV 10

RESULT 8  
US-09-001-984C-97  
Sequence 97, Application US/09001984C  
Patent No. 6245331  
GENERAL INFORMATION:  
APPLICANT: Laal, Susan  
APPLICANT: Zolla-Pazner, Susan  
TITLE OF INVENTION: EARLY DETECTION OF MYCOBACTERIAL DISEASE  
FILE REFERENCE: NYU-011  
CURRENT APPLICATION NUMBER: US/09/001,984C  
CURRENT FILING DATE: 1997-12-31  
PRIOR APPLICATION NUMBER: 60/034,003  
PRIOR FILING DATE: 1996-12-31  
NUMBER OF SEQ ID NOS: 106  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 97  
LENGTH: 10  
TYPE: PRT

ORGANISM: Mycobacterium tuberculosis strain H37RV  
US-09-001-984C-97

Query Match 50.0%; Score 26; DB 3; Length 10;  
Best Local Similarity 80.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIRY 6  
||:|  
DB 6 LPEY 10

RESULT 9  
US-09-001-984C-101

; Sequence 101, Application US/09001984C  
; Patent No. 6245331  
; GENERAL INFORMATION:  
; APPLICANT: Laal, Suman  
; APPLICANT: Zolla-Pazner, Susan  
; APPLICANT: Belisle, John T  
; TITLE OF INVENTION: EARLY DETECTION OF MYCOBACTERIAL DISEASE  
; FILE REFERENCE: NYU-011  
; CURRENT APPLICATION NUMBER: US/09/001,984C  
; CURRENT FILING DATE: 1997-12-31  
; PRIOR APPLICATION NUMBER: 60/034,003  
; PRIOR FILING DATE: 1996-12-31  
; NUMBER OF SEQ ID NOS: 106  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 101  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Mycobacterium tuberculosis strain H37RV  
US-09-001-984C-101

Query Match 50.0%; Score 26; DB 3; Length 10;  
Best Local Similarity 80.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIRY 6  
||:|  
DB 6 LPEY 10

RESULT 10  
US-09-396-347F-92  
; Sequence 92, Application US/09396347F  
; Patent No. 6506384  
; GENERAL INFORMATION:  
; APPLICANT: Laal, Suman  
; APPLICANT: Zolla-Pazner, Susan  
; APPLICANT: Belisle, John T  
; TITLE OF INVENTION: EARLY DETECTION OF MYCOBACTERIAL DISEASE  
; FILE REFERENCE: 32004-169276  
; CURRENT APPLICATION NUMBER: US/09/396,347F  
; CURRENT FILING DATE: 1999-09-14  
; PRIOR APPLICATION NUMBER: 09/001,984  
; PRIOR FILING DATE: 1997-12-31  
; NUMBER OF SEQ ID NOS: 106  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 92  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Mycobacterium tuberculosis strain H37RV  
US-09-396-347F-92

Query Match 50.0%; Score 26; DB 4; Length 10;  
Best Local Similarity 80.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIRY 6  
||:|  
DB 6 LPEY 10

RESULT 11  
US-09-396-347F-97

; Sequence 97, Application US/09396347F  
; Patent No. 6506384  
; GENERAL INFORMATION:  
; APPLICANT: Laal, Suman  
; APPLICANT: Zolla-Pazner, Susan  
; APPLICANT: Belisle, John T  
; TITLE OF INVENTION: EARLY DETECTION OF MYCOBACTERIAL DISEASE  
; FILE REFERENCE: 32004-169276  
; CURRENT APPLICATION NUMBER: US/09/396,347F  
; CURRENT FILING DATE: 1999-09-14  
; PRIOR APPLICATION NUMBER: 09/001,984  
; PRIOR FILING DATE: 1997-12-31  
; NUMBER OF SEQ ID NOS: 106  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 97  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Mycobacterium tuberculosis strain H37RV  
US-09-396-347F-97

Query Match 50.0%; Score 26; DB 4; Length 10;  
Best Local Similarity 80.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIRY 6  
||:|  
DB 6 LPEY 10

RESULT 12  
US-09-396-347F-101  
; Sequence 101, Application US/09396347F  
; Patent No. 6506384  
; GENERAL INFORMATION:  
; APPLICANT: Laal, Suman  
; APPLICANT: Zolla-Pazner, Susan  
; APPLICANT: Belisle, John T  
; TITLE OF INVENTION: EARLY DETECTION OF MYCOBACTERIAL DISEASE  
; FILE REFERENCE: 32004-169276  
; CURRENT APPLICATION NUMBER: US/09/396,347F  
; CURRENT FILING DATE: 1999-09-14  
; PRIOR APPLICATION NUMBER: 09/001,984  
; PRIOR FILING DATE: 1997-12-31  
; NUMBER OF SEQ ID NOS: 106  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 101  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Mycobacterium tuberculosis strain H37RV  
US-09-396-347F-101

Query Match 50.0%; Score 26; DB 4; Length 10;  
Best Local Similarity 80.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPIRY 6  
||:|  
DB 6 LPEY 10

RESULT 13  
US-08-604-965E-10

; Sequence 10, Application US/08604965E  
; Patent No. 6046033  
; GENERAL INFORMATION:  
; APPLICANT: Goto, Masaki  
; APPLICANT: Tsuda, Sioke  
; APPLICANT: Yano, Kazuki  
; APPLICANT: Kobayashi, Fumie  
; APPLICANT: Yamaguchi, Kyoji

APPLICANT: Washida, Naohiro  
APPLICANT: Satake, Toshiko  
APPLICANT: Morinaga, Tomonori  
APPLICANT: Ueda, Masatsugu  
APPLICANT: Higashio, Kanji  
TITLE OF INVENTION: Basic Osteoblast Growth Factor II (bOGF)  
TITLE OF INVENTION: II  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burgess, Ryan and Wayne  
STREET: 370 Lexington Avenue, Suite 2105  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 10017  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5 1/4 inch diskette  
COMPUTER: PC'S LIMITED SYSTEM 200  
OPERATING SYSTEM: DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/604,965E  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP95/01270  
FILING DATE: June 27, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Wayne, Milton J.  
REGISTRATION NUMBER: 17,906  
REFERENCE/DOCKET NUMBER: U-WP-5212  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-683-8150  
TELEFAX: 212-532-4285  
TELEX: 423794  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-604-965E-10

Query Match 48.1%; Score 25; DB 3; Length 7;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 EYGP 8  
|||  
Db 3 EYGP 6

RESULT 14  
US-08-604-965E-15  
Sequence 15, Application US/08604965E  
Patent No. 6046033  
GENERAL INFORMATION:  
APPLICANT: Goto, Masaaki  
APPLICANT: Tsuda, Eisuke  
APPLICANT: Yano, Kazuki  
APPLICANT: Kobayashi, Fumie  
APPLICANT: Yamaguchi, Kyoji  
APPLICANT: Washida, Naohiro  
APPLICANT: Satake, Toshiko  
APPLICANT: Morinaga, Tomonori  
APPLICANT: Ueda, Masatsugu  
APPLICANT: Higashio, Kanji  
TITLE OF INVENTION: Basic Osteoblast Growth Factor II (bOGF)  
TITLE OF INVENTION: II  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burgess, Ryan and Wayne

STREET: 370 Lexington Avenue, Suite 2105  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 10017  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5 1/4 inch diskette  
COMPUTER: PC'S LIMITED SYSTEM 200  
OPERATING SYSTEM: DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/604,965E  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP95/01270  
FILING DATE: June 27, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Wayne, Milton J.  
REGISTRATION NUMBER: 17,906  
REFERENCE/DOCKET NUMBER: U-WP-5212  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-683-8150  
TELEFAX: 212-532-4285  
TELEX: 423794  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-604-965E-15

Query Match 48.1%; Score 25; DB 3; Length 7;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 EYGP 8  
|||  
Db 3 EYGP 6

RESULT 15  
5258287-43  
Patent No. 5258287  
APPLICANT: BAXTER, ROBERT C.; WOOD, WILLIAM I.  
TITLE OF INVENTION: DNA ENCODING AND METHODS OF PRODUCTION  
OF INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN BP53  
NUMBER OF SEQUENCES: 58  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/171,623  
FILING DATE: 22-MAR-1988  
SEQ ID NO: 43  
LENGTH: 7  
5258287-43

Query Match 48.1%; Score 25; DB 6; Length 7;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 EYGP 8  
|||  
Db 3 EYGP 6

Search completed: November 30, 2004, 08:15:59  
Job time : 28 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:41:32, Search time 11.6667 Seconds  
(without alignments)  
82.472 Million cell updates/sec

Title: US-10-008-377a-2  
Perfect score: 52  
Sequence: 1 GVLMWELFSL 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 1102

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database: PIR 791\*  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	46.2	10	2	E49033 T-cell receptor ga
2	20	38.5	10	2	F49033 T-cell receptor ga
3	19	36.5	9	2	D57444 neurotrophin Grb-A
4	18	34.6	9	2	I58350 gene C-mpl protein
5	17	32.7	6	2	PT0519 T-cell receptor be
6	17	32.7	7	2	PT0008 glucuronosyltransf
7	17	32.7	10	2	S71868 glutathione transf
8	17	32.7	10	2	S63696 DNA polymerase - Y
9	16	30.8	5	2	PT0308 Ig heavy chain CRD
10	16	30.8	6	2	S66195 alcohol dehydrogen
11	15	30.8	6	2	B34835 dhaa protein - Pse
12	15	30.8	6	2	B31263 dihydrofolate redu
13	15	30.8	6	2	B31263 dihydrofolate redu
14	15	30.8	6	2	B35640 cerebellar degener
15	16	30.8	7	2	S09652 hypochlorite prote
16	16	30.8	8	2	A59495 vesicle associated
17	16	30.8	8	2	A59495 vesicle associated
18	16	30.8	8	2	C61512 variant surface gl
19	16	30.8	8	2	D61512 variant surface gl
20	16	30.8	8	2	A41117 acetylcholinestera
21	16	30.8	10	2	S59625 beta-galactosidase
22	16	30.8	10	2	PT0310 Ig heavy chain CRD
23	16	30.8	10	2	S66214 cartilage oligomer
24	15	28.8	7	2	PH1602 Ig H chain V-D-U r
25	15	28.8	7	2	S13567 tubulin beta-3 cha
26	15	28.8	8	2	S13288 acylase - Kluverera
27	15	28.8	9	2	S07241 litorin - Rohde's
28	14	26.9	5	2	PT0278 Ig heavy chain CRD
29	14	26.9	5	2	S69237 surface protein te

30	14	26.9	7	2	S33244 neuromodulatory pe
31	14	26.9	7	2	S33246 neuromodulatory pe
32	14	26.9	9	2	A61386 macrophage inhibic
33	14	26.9	10	2	A60624 angiotensin I - Ja
34	14	26.9	10	2	B38887 T-cell receptor ga
35	14	26.9	10	2	S05964 hypothetical prote
36	14	26.9	10	2	A99917 angiotensin precu
37	14	26.9	10	2	T17075 cyclochrome-c oxida
38	14	26.9	10	2	G60787 sperm-activating p
39	14	26.9	10	2	E60787 sperm-activating p
40	14	26.9	10	2	C60787 sperm-activating p
41	14	26.9	10	2	C60588 sperm-activating p
42	14	26.9	10	2	A60588 sperm-activating p
43	14	26.9	10	2	E60527 sperm-activating p
44	14	26.9	10	2	E39572 sperm-activating p
45	14	26.9	10	2	D60589 sperm-activating p

## ALIGNMENTS

## RESULT 1

E49033 T-cell receptor gamma chain V-D-J region - human (fragment)  
C.Species: Homo sapiens (man)  
C.Date: 19-Dec-1993 #sequence\_revision 25-Aug-1995 #text\_change 21-Jul-2000  
C.Accession: E49033; D49033  
R.Morita, C.T.; Verma, S.; Aparicio, P.; Martinez, C.; Spits, H.; Brenner, M.B.  
Eur. J. Immunol. 21, 2999-3007, 1991  
A>Title: Functionally distinct subsets of human gamma/delta T cells.  
A.Reference number: A49033; PMID:92083926; PMID:1684157  
A.Accession: E49033  
A.Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-10 <MOR>  
A.Cross-references: GB:S72587; NID:G240696; PIDN:AA520630.1; PID:G240697  
A.Note: sequence extracted from NCBI backbone (NCBIN:72591, NCBI:P:72595)  
A.Accession: D49033  
A.Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-10 <MOR>  
A.Cross-references: GB:S72587; NID:G240696; PIDN:AA520630.1; PID:G240697  
A.Note: sequence extracted from NCBI backbone (NCBIN:72587, NCBI:P:72589)  
C.Keywords: T-cell receptor

Query Match 46.2%; Score 24; DB 2; Length 10;  
Best local similarity 57.1%; Pred. No. 2.6e+02;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 4 LMEWELFSL 10  
DB 3 LMEWELFSL 9

## RESULT 2

E49033 T-cell receptor gamma chain V-D-J region - human (fragment)  
C.Species: Homo sapiens (man)  
C.Date: 19-Dec-1993 #sequence\_revision 17-Mar-2000 #text\_change 17-Mar-2000  
C.Accession: E49033  
R.Morita, C.T.; Verma, S.; Aparicio, P.; Martinez, C.; Spits, H.; Brenner, M.B.  
Eur. J. Immunol. 21, 2999-3007, 1991  
A>Title: Functionally distinct subsets of human gamma/delta T cells.  
A.Reference number: A49033; PMID:92083926; PMID:1684157  
A.Accession: E49033  
A.Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-10 <MOR>  
A.Cross-references: GB:S72605; NID:G240700; PIDN:AA520632.1; PID:G240701  
A.Note: sequence extracted from NCBI backbone (NCBIN:72605, NCBI:P:72606)  
C.Keywords: T-cell receptor

Query Match 38.5%; Score 20; DB 2; Length 10;

Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LME 6  
|||  
Db 3 LME 5

## RESULT 3

D57444

neuropeptide Grp-Asp B4 - two-spotted cricket

C/Species: Gryllus bimaculatus (two-spotted cricket)

C/Date: 26-Jan-1996 #sequence\_revision 26-Jan-1996 #ext\_change 09-Jul-2004

C/Accession: D57444

R/Lorenz, M.W.; Kellner, R.; Hoffmann, K.H.

J. Biol. Chem. 270: 21103-21108, 1995

A/Title: A family of neuropeptides that inhibit juvenile hormone biosynthesis in the cr

A/Reference number: A57444; MUID:95403341; PMID:7673141

A/Accession: D57444

A/Status: preliminary

A/Molecule type: protein

A/Residues: 1-9 &lt;LOR&gt;

A/Cross-references: UNIPROT:Q7M3N6

## Query Match

Best Local Similarity 36.5%; Score 19; DB 2; Length 9;

Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 WEIF 8  
|||  
Db 2 WEIF 5

## RESULT 4

I58350

gene c-mpl protein - mouse (fragment)

C/Species: Mus sp. (mouse)

C/Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #ext\_change 28-Feb-1997

C/Accession: I58350

R/Alexander, W.S.; Dunn, A.R.

Oncogene 10, 795-803, 1995

A/Title: Structure and transcription of the genomic locus encoding murine c-mpl, a recep

A/Reference number: I58350; MUID:95166571; PMID:7862460

A/Accession: I58350

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-9 &lt;RMS&gt;

A/Cross-references: GB:S76841; NID:9312992

C/Genetics:

A/Accession: C-mpl

## Query Match

Best Local Similarity 34.6%; Score 18; DB 2; Length 9;

Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 WEIF 8  
|||  
Db 4 WAF 7

## RESULT 5

PT0519

T-cell receptor beta chain V-D-J region (100-4C) - mouse (fragment)

C/Species: Mus musculus (house mouse)

C/Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #ext\_change 30-May-1997

C/Accession: PT0519

R/Peeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A/Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A/Reference number: PT0509; MUID:91277601; PMID:1711558

A/Accession: PT0519

A/Status: translation not shown

A/Molecule type: mRNA

A/Residues: 1-6 <PEZ>  
A/Experimental source: adult thymus, strain BALB/c  
C/Keywords: T-cell receptor

Query Match 32.7%; Score 17; DB 2; Length 6;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 LME 6  
|||  
Db 4 LMD 6

## RESULT 6

PX0008

glucuronosyltransferase (EC 2.4.1.17), hepatic - rat (fragment)

N/Alternate names: UDP-glucuronosyltransferase

C/Species: Rattus norvegicus (Norway rat)

C/Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #ext\_change 07-Feb-1997

C/Accession: PX0008

R/Yokota, H.; Yuasa, A.; Sato, R.

J. Biochem. 104, 531-536, 1988

A/Title: Purification and properties of a form of UDP-glucuronosyltransferase from liver "

A/Reference number: PX0008; MUID:89197852; PMID:3149280

A/Accession: PX0008

A/Molecule type: protein

A/Residues: 1-7 &lt;YOK&gt;

A/Keywords: glucosyltransferase, hexosyltransferase, liver

Query Match 32.7%; Score 17; DB 2; Length 7;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLM 5  
|||  
Db 3 LVM 6

## RESULT 7

S71868

N/A/Alternate names: glutathione S-transferase class mu 4 - pig (fragment)

C/Species: Sus scrofa domestica (domestic pig)

C/Date: 19-Mar-1998 #sequence\_revision 13-Sep-1998 #ext\_change 09-Jul-2004

C/Accession: S71868

R/Rountz, P.; Anglade, P.; Debrauer, L.; Tulliez, J.

Biochem. J. 317, 879-884, 1998

A/Title: Characterization of pig liver glutathione S-transferases using HPLC-electrospr

A/Reference number: S71864; MUID:96332484; PMID:9760377

A/Accession: S71868

A/Molecule type: protein

A/Residues: 1-10 &lt;ROU&gt;

A/Cross-references: UNIPROT:Q7M3B8

C/Comment: At least five species-independent classes of cytosolic glutathion transferas

s mitochondrial form are known.

C/Complex: dimer

C/Function:

A/Description: catalyzes the nucleophilic conjugation of intracellular glutathione to a

A/Pathway: detoxification; xenobiotics metabolism

A/Note: increased hydrophilicity of GSH-conjugates facilitates their further metabolism

C/Superfamily: glutathione transferase

C/Keywords: dimer; transferase

Query Match 32.7%; Score 17; DB 2; Length 10;  
Best Local Similarity 66.7%; Pred. No. 4.2e+03;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 WEI 7  
|||  
Db 6 WDI 8

RESULT 8  
 S63696  
 DNA polymerase - yeast (*Kluyveromyces marxianus* var. *lactis*) killer plasmid PGKL2 (fragment)  
 C:Species: *Kluyveromyces marxianus* var. *lactis*, *Candida sphacelata*  
 C:Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 25-Apr-1997  
 C:Accession: S63696  
 R:Raked, M.; Hiraiishi, H.; Takesako, T.; Tanase, S.; Gunze, N.  
 Year: 12, 241-246, 1996  
 A:Title: The terminal protein of the linear DNA plasmid PGKL2 shares an N-terminal domain  
 A:Reference number: S63696; PMID:97060015; PMID:8904336  
 A:Accession: S63696  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-10 <TAK>  
 A>Note: the source is designated as *Kluyveromyces lactis*

Query Match  
 Best Local Similarity 32.7%; Score 17; DB 2; Length 10;  
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 VLLME 6  
 DB 1 VXXME 5

RESULT 9  
 PT0308  
 Ig heavy chain CRD3 region (clone 6-88) - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
 C:Accession: PT0308  
 R:Yanada, M.; Maserman, R.; Reichard, B.A.; Shane, S.; Catton, A.J.; Rovera, G.  
 J. Exp. Med. 173, 395-407, 1991  
 A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and  
 A:Reference number: PT0222; PMID:91108337; PMID:1899102  
 A:Accession: PT0308  
 A:Molecule type: DNA  
 A:Residues: 1-5 <YAM>  
 A:Experimental source: B lymphocyte  
 C:Keywords: heterotrimer; immunoglobulin

Query Match  
 Best Local Similarity 30.8%; Score 16; DB 2; Length 5;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WE 6  
 DB 2 WE 3

RESULT 10  
 S66195  
 alcohol dehydrogenase (EC 1.1.1.1) class III low affinity form - cod (Gadus sp.) (fragment)  
 C:Species: *Gadus* sp. (cod)  
 C:Date: 19-Mar-1997 #sequence\_revision 29-Aug-1997 #text\_change 12-Jun-1998  
 C:Accession: S66195  
 R:Hilmy, L.; Hackett, M.; Shaqat, J.; Danielsson, O.; Iida, J.; Hendrickson, R.C.;  
 FEBS Lett. 367, 237-240, 1995  
 A:Title: Multiplicity of N-terminal structures of medium-chain alcohol dehydrogenases. M  
 nzyme.  
 A:Reference number: S66191; PMID:95331382; PMID:7607314  
 A:Accession: S66195  
 A:Molecule type: protein  
 A:Residues: 1-6 <HJE>  
 C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology  
 C:Keywords: alcohol metabolism; NAD; oxidoreductase

Query Match  
 Best Local Similarity 30.8%; Score 16; DB 2; Length 6;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WE 6  
 DB 1

DB 5 WE 6  
 RESULT 11  
 B34835  
 dhaA protein - *Pseudomonas aeruginosa* (fragment)  
 C:Species: *Pseudomonas aeruginosa*  
 C:Date: 13-Jul-1990 #sequence\_revision 13-Jul-1990 #text\_change 08-Oct-1999  
 C:Accession: B34835  
 R:Yee, T.W.; Smith, D.W.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 1278-1282, 1990  
 A:Title: Pseudomonas chromosomal replication origins: a bacterial class distinct from E  
 A:Reference number: A34835; PMID:90160310; PMID:2106132  
 A:Accession: B34835  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-6 <YEE>  
 A:Cross-references: GB:M30125; NID:9151419; PID:AAA25916.1; PID:9151421  
 C:Keywords: DNA binding

Query Match  
 Best Local Similarity 30.8%; Score 16; DB 2; Length 6;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLLM 5  
 DB 3 VELL 6

RESULT 12  
 A31263  
 dihydrofolate reductase (EC 1.5.1.3) / thymidylate synthase (EC 2.1.1.45) - *Plasmodium falciparum*  
 C:Species: *Plasmodium falciparum*  
 C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 20-Mar-1996  
 C:Accession: A31263  
 R:Peterson, D.S.; Walliker, D.; Wellens, T.E.  
 Proc. Natl. Acad. Sci. U.S.A. 85, 9114-9118, 1988  
 A:Title: Evidence that a point mutation in dihydrofolate reductase-thymidylate synthase  
 A:Reference number: A94217; PMID:89057886; PMID:2904149  
 A:Accession: A31263  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-6 <PET>  
 C:Keywords: methyltransferase; NADP; oxidoreductase

Query Match  
 Best Local Similarity 30.8%; Score 16; DB 2; Length 6;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WE 6  
 DB 3 WE 4

RESULT 13  
 B31253  
 dihydrofolate reductase (EC 1.5.1.3) / thymidylate synthase (EC 2.1.1.45) - *Plasmodium falciparum*  
 C:Species: *Plasmodium falciparum*  
 C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 20-Mar-1996  
 C:Accession: B31253  
 R:Peterson, D.S.; Walliker, D.; Wellens, T.E.  
 Proc. Natl. Acad. Sci. U.S.A. 85, 9114-9118, 1988  
 A:Title: Evidence that a point mutation in dihydrofolate reductase-thymidylate synthase  
 A:Reference number: A94217; PMID:89057886; PMID:2904149  
 A:Accession: B31253  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-6 <PET>  
 C:Keywords: methyltransferase; NADP; oxidoreductase

Query Match  
 Best Local Similarity 30.8%; Score 16; DB 2; Length 6;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WE 6  
||  
Db 3 WE 4

## RESULT 14

B35640  
cerebellar degeneration-related protein - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 28-Sep-1990 #sequence\_revision 28-Sep-1990 #text\_change 24-Jun-1993  
C:Accession: B35640  
R:Chen, Y.T.; Rettig, W.J.; Yemamandra, A.K.; Kozak, C.A.; Chaganti, R.S.K.; Posner, J.E  
Proc. Natl. Acad. Sci. U.S.A. 87, 3077-3081, 1990  
A:Title: Cerebellar degeneration-related antigen: a highly conserved neuroectodermal marker  
A:Reference number: A35640; MUID:90222173; PMID:2326268  
A:Accession: B35640  
A>Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-6 <CHE>

Query Match 30.8%; Score 16; DB 2; Length 6;  
Best local similarity 100.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 0; Gaps 0;

QY 5 WE 6  
||  
Db 2 WE 3

## RESULT 15

S09652  
hypothetical protein (aacC2 3' region) - Enterobacter cloacae (fragment)  
C:Species: Enterobacter cloacae  
C:Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 08-Oct-1999  
C:Accession: S09652  
R:Wiegand, J.S.; Kerebelar-van Gaalen, P.A.G.; van de Klundert, J.A.M.  
Antimicrob. Agents Chemother. 33, 1153-1159, 1989  
A:Title: Nucleotide sequence of the aacC2 gene, a gentamicin resistance determinant invc  
A:Reference number: S09651; MUID:90024972; PMID:2552900  
A:Accession: S09652  
A:Molecule type: DNA  
A:Residues: 1-7 <VLI>  
A:Cross-references: EMBL:X51534; NID:940879; PID:CAA35914.1; PID:9581034

Query Match 30.8%; Score 16; DB 2; Length 7;  
Best local similarity 25.0%; Pred. No. 2.8e+05;  
Matches 1; Conservative 3; Mismatches 0; Gaps 0;

QY 2 VLLW 5  
:::  
Db 1 MLTW 4

Search completed: November 30, 2004, 08:14:26  
Job time : 12.6667 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:40:12 / Search time 78.3333 Seconds  
(without alignments)  
73.452 Million cell updates/sec

Title: US-10-008-377A-2  
Perfect score: 52  
Sequence: 1 GVLWEIFSL 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1825161 seqs, 57537466 residues

Total number of hits satisfying chosen parameters: 2971

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: UniProt\_02: \*  
1: uniProt\_sprot: \*  
2: uniProt\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	44.2	10	2	06JL97	06JL97	neisseria g
2	44.2	10	2	AA516521	AA516521	neisseria g
3	42.3	10	2	Q8JFE7	Q8JFE7	flicicula al
4	42.3	10	2	Q8JFE7	Q8JFE7	flicicula al
5	38.5	10	2	Q8JFE7	Q8JFE7	flicicula al
6	38.5	10	2	Q8JFE7	Q8JFE7	flicicula al
7	36.5	10	2	Q8JFE7	Q8JFE7	flicicula al
8	36.5	10	2	Q8JFE7	Q8JFE7	flicicula al
9	36.5	10	2	Q8JFE7	Q8JFE7	flicicula al
10	36.5	10	2	Q8JFE7	Q8JFE7	flicicula al
11	34.6	10	2	Q8JFE7	Q8JFE7	flicicula al
12	34.6	10	2	Q8JFE7	Q8JFE7	flicicula al
13	34.6	10	2	Q8JFE7	Q8JFE7	flicicula al
14	32.7	10	2	Q8JFE7	Q8JFE7	flicicula al
15	32.7	10	2	Q8JFE7	Q8JFE7	flicicula al
16	32.7	10	2	Q8JFE7	Q8JFE7	flicicula al
17	32.7	10	2	Q8JFE7	Q8JFE7	flicicula al
18	32.7	10	2	Q8JFE7	Q8JFE7	flicicula al
19	32.7	10	2	Q8JFE7	Q8JFE7	flicicula al
20	32.7	10	2	Q8JFE7	Q8JFE7	flicicula al
21	32.7	10	2	Q8JFE7	Q8JFE7	flicicula al
22	32.7	10	2	Q8JFE7	Q8JFE7	flicicula al
23	32.7	10	2	Q8JFE7	Q8JFE7	flicicula al
24	32.7	10	2	Q8JFE7	Q8JFE7	flicicula al
25	30.8	10	2	Q8JFE7	Q8JFE7	flicicula al
26	30.8	10	2	Q8JFE7	Q8JFE7	flicicula al
27	30.8	10	2	Q8JFE7	Q8JFE7	flicicula al
28	30.8	10	2	Q8JFE7	Q8JFE7	flicicula al
29	30.8	10	2	Q8JFE7	Q8JFE7	flicicula al
30	30.8	10	2	Q8JFE7	Q8JFE7	flicicula al
31	30.8	10	2	Q8JFE7	Q8JFE7	flicicula al

32	15	30.8	9	2	070YA3	070YA3	aeollanthus
33	16	30.8	9	2	09AYG2	09AYG2	varanus ind
34	16	30.8	9	2	085723	085723	simlan sarc
35	16	30.8	9	2	CAD45451	CAD45451	aeollanth
36	16	30.8	10	2	Q7SA62	Q7SA62	neurospora
37	16	30.8	10	2	Q8SHB4	Q8SHB4	furcifer ve
38	16	30.8	10	2	Q8SHB7	Q8SHB7	furcifer ou
39	16	30.8	10	2	Q8SHC0	Q8SHC0	furcifer la
40	16	30.8	10	2	Q8SHC3	Q8SHC3	furcifer la
41	15.5	29.8	10	2	Q8BAX0	Q8BAX0	notophthalm
42	15.5	29.8	10	2	Q8SHD2	Q8SHD2	chamaeleo w
43	15.5	29.8	10	2	Q8SHD5	Q8SHD5	chamaeleo s
44	15.5	29.8	10	2	Q8SHD8	Q8SHD8	chamaeleo r
45	15.5	29.8	10	2	Q8SHD4	Q8SHD4	chamaeleo g

## ALIGNMENTS

```

RESULT 1
ID Q6JL97
AC Q6JL97
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE Nucl (Fragment).
GN Name=nucl.
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MS11;
RX PubMed=15084227;
RA Snyder L.A., Davies J.K., Saunders N.J.;
RT "Microarray genotyping of key experimental strains of Neisseria
RT gonorrhoeae reveals gene complement diversity and five new neisserial
RT genes associated with Minimal Mobile Elements."
RL BMC Genomics 5:23-23(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MS11;
RA Snyder L.A.S., Davies J.K., Saunders N.J.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY386266; AA516521.1; -
FT NON_TER
SQ
SEQUENCE 10 AA; 1227 MW; BACCB286379D1A6 CRC64;

Query Match 44.2%; Score 23; DB 2; Length 10;
Best Local Similarity 37.5%; Pred. No. 2.6e+03;
Matches 3; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GVLWEIF 8
DB 2 GMTFWGLF 9

RESULT 2
ID AA516521
AC AA516521
DT 20-MAY-2004 (TREMBLrel. 27, Created)
DT 20-MAY-2004 (TREMBLrel. 27, Last sequence update)
DE Nucl (Fragment).
GN Nucl.
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=MS11;
RX PubMed=15084227;
RA Snyder L.A., Davies J.K., Saunders N.J.;
RT "Microarray genotyping of key experimental strains of Neisseria
RT gonorrhoeae reveals gene complement diversity and five new neisseria
RT genes associated with Minimal Mobile Elements.";
RL BMC Genomics 5:23-23(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MS11;
RA Snyder L.A.S., Davies J.K., Saunders N.J.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY386266; AAS16521.1; -.
FT NON_TER
SQ SEQUENCE 10 AA; 1227 MW; BACCB386379D1A6 CRC64;

Query Match
Best Local Similarity 44.2%; Score 23; DB 2; Length 10;
Matches 3; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GVLWLF 8
DB 2 GMTFWGLF 9

RESULT 3
ID Q8UF7 PRELIMINARY; PRT; 10 AA.
AC Q8UF7;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 22, Last annotation update)
DE Myelin proteolipid protein (Fragment).
OS Picea alba (Collected flycatcher).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinoptera; Aves; Neognathae; Passeriformes; Muscicapidae; Ficedula.
OC NCBI_TaxID=59894;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21918460; PubMed=11918793;
RA Primer C.R., Borge T., Lindell U., Saetre G.P.;
RT "Single-nucleotide polymorphism characterization in species with
RT limited available sequence information: high nucleotide diversity
RT revealed in the avian genome.";
RL Mol. Ecol. 11:603-612(2002).
DR EMBL; AF454217; AAM22903.1; -.
FT NON_TER
SQ SEQUENCE 10 AA; 1206 MW; 1FAAC96732C86B CRC64;

Query Match
Best Local Similarity 42.3%; Score 22; DB 2; Length 10;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVLW 5
DB 4 GVLW 8

RESULT 4
ID Q8UJ3 PRELIMINARY; PRT; 10 AA.
AC Q8UJ3;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Myelin proteolipid protein (Fragment).
OS Ficedula hypoleuca.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinoptera; Aves; Neognathae; Passeriformes; Muscicapidae; Ficedula.
OC NCBI_TaxID=46689;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21918460; PubMed=11918793;
RA Primer C.R., Borge T., Lindell U., Saetre G.P.;
RT "Single-nucleotide polymorphism characterization in species with
RT limited available sequence information: high nucleotide diversity
RT revealed in the avian genome.";
RL Mol. Ecol. 11:603-612(2002).
DR EMBL; AF454216; AAM22902.1; -.
FT NON_TER
SQ SEQUENCE 10 AA; 1206 MW; 1FAAC96732C86B CRC64;

Query Match
Best Local Similarity 42.3%; Score 22; DB 2; Length 10;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVLW 5
DB 4 GVLW 8

RESULT 5
ID Q85406 PRELIMINARY; PRT; 8 AA.
AC Q85406;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Coxiella burnetii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Coxiellaceae; Coxiella.
OC NCBI_TaxID=777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nine Mile Phase I;
RX MEDLINE=98348442; PubMed=9683477;
RA Williams H., Jaeger C., Baljer G.;
RT "Physical and genetic map of the obligate intracellular bacterium
RT Coxiella burnetii.";
RL J. Bacteriol. 180:3816-3822(1998).
DR EMBL; AF064963; AAD09947.1; -.
FT NON_TER
SQ SEQUENCE 8 AA; 993 MW; 046BSAA53772727 CRC64;

Query Match
Best Local Similarity 38.5%; Score 20; DB 2; Length 8;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLLW 5
DB 1 LLLW 4

RESULT 6
ID Q8SHF6 PRELIMINARY; PRT; 10 AA.
AC Q8SHF6;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
OS Chamaeleo melleri (Meller's chameleon).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodactylidae; Squamata; Iguania; Acrodonta; Chamaeleonidae; Chamaeleo.
OC NCBI_TaxID=179915;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22169767; PubMed=12162400;

```

RA Townsend T., Larson A.;  
 RT "Molecular phylogenetics and mitochondrial genomic evolution in the  
 RT chameleons (Reptilia, Squamata).";  
 RL Mol. Phylogenet. Evol. 23:22-36(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Townsend T.M., Larson A.L.;  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF448753; AAL90547.1; -  
 DR GO; GO:0005739; Mitochondrion; IEA.  
 KM Mitochondrion.  
 FT NON TER  
 SQ SEQUENCE 10 AA; 1276 MW; 5B21B2733772727 CRC64;

Query Match 38.5%; Score 20; DB 2; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 9.1e+03;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLW 5  
 Db 3 LILW 6

## RESULT 7

Q8JB81 PRELIMINARY; PRT; 7 AA.  
 AC Q8JB81;  
 DT 01-OCT-2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
 DE Truncated pol protein (Fragment).  
 GN Name-pol;  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22060770;  
 RA Beerenwinkel N., Schmidt B., Walter H., Kaiser R., Lengauer T.,  
 RA Hoffmann D., Korn K., Selbig U.;  
 RT "Diversity and complexity of HIV-1 drug resistance: a bioinformatics  
 RT approach to predicting phenotype from genotype.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:8271-8276(2002).  
 DR EMBL; AF347267; AAK32344.1; -  
 FT NON TER  
 SQ SEQUENCE 7 AA; 885 MW; 76C37731A046C700 CRC64;

Query Match 36.5%; Score 19; DB 2; Length 7;  
 Best Local Similarity 40.0%; Pred. No. 1.8e+06;  
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLW 6  
 Db 3 LILW 7

## RESULT 8

Q7M3N6 PRELIMINARY; PRT; 9 AA.  
 AC Q7M3N6;  
 DT 01-MAR-2004 (TREMBlrel. 26, Created)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Neuropeptide Grb-AST B4.  
 OS Gylilus bimaculatus (Two-spotted cricket).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preygotia;  
 OC Neoptera; Orthopteroidea; Orthoptera; Ensifera; Gryllidae; Gryllinae;  
 OC Gryllus.  
 OX NCBI\_TaxID=6999;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95403341; PubMed=7673141;  
 RA Lorenz M.W., Kellner R., Hoffmann K.H.;

RT "A family of neuropeptides that inhibit juvenile hormone biosynthesis  
 RT in the cricket, Gylilus bimaculatus.";  
 RL J. Biol. Chem. 270:21103-21108(1995).  
 DR PIR; D57444; D57444.  
 SQ SEQUENCE 9 AA; 1175 MW; 3860B871E9D40B03 CRC64;

Query Match 36.5%; Score 19; DB 2; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 WEIF 8  
 Db 2 WEFF 5

## RESULT 9

Q8WTT4 PRELIMINARY; PRT; 10 AA.  
 AC Q8WTT4;  
 DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Dyscrophin (Fragment).  
 GN Name=dystrophin; Synonyms=DMD;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ditea S.D., Klamut H.J., Ray P.N., Worton R.G.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Beggs A.H., Koenig M., Kunkel L.M.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA den Dunnen J.T.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF276053; AAL35752.1; -  
 DR EMBL; AF213401; AAL61549.1; -  
 FT NON TER  
 SQ SEQUENCE 10 AA; 1340 MW; 2B999202CB1B0363 CRC64;

Query Match 36.5%; Score 19; DB 2; Length 10;  
 Best Local Similarity 60.0%; Pred. No. 1.4e+04;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLW 6  
 Db 1 MLWWE 5

## RESULT 10

Q8OX07 PRELIMINARY; PRT; 9 AA.  
 AC Q8OX07;  
 DT 01-JUN-2003 (TREMBlrel. 24, Created)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Thrombopoietin receptor (Fragment).  
 GN Name=c-mpl;  
 OS Mus sp.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10095;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9516571; PubMed=7862460;  
 RA Alexander W.S., Dunn A.R.;  
 RT "Structure and transcription of the genomic locus encoding murine c-  
 RT Mpl, a receptor for thrombopoietin.";



Query Match 32.7%; Score 17; DB 2; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1.8e+06;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0;

QY 4 LME 6  
 11:  
 Db 6 LMD 8

## RESULT 15

Q25355 PRELIMINARY; PRT; 10 AA.  
 AC Q25355; 01-NOV-1996 (TREMREL.01, Created)  
 DT 01-NOV-1996 (TREMREL.01, Last sequence update)  
 DT 01-NOV-1998 (TREMREL.08, Last annotation update)  
 DE Viteillogenin A (Fragment).  
 OS Locusta migratoria (Migratory locust).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridoidea;  
 OC Acridoidea; Acrididae; Oedipodinae; Locusta.  
 OX NCBI\_Taxid=7004;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=8603971; PubMed=2820677;  
 RA Locke J., White B.N., Wyatt G.R.;  
 RT "Cloning and 5' end nucleotide sequences of two juvenile hormone-  
 inducible vitellogenin genes of the African migratory locust.",  
 RL DNA 6:331-342(1987).  
 DR EMBL; M17333; AAA29284.1; -.  
 FT NON TER 10  
 SQ SEQUENCE 10 AA; 1128 MW; D1B31177272042CD CRC64;

Query Match 32.7%; Score 17; DB 2; Length 10;  
 Best Local Similarity 28.6%; Pred. No. 3.2e+04;  
 Matches 2; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 LMEFSL 10  
 11:  
 Db 1 MMAYILL 7

Search completed: November 30, 2004, 08:13:40  
 Job time : 79.333 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:41:57; Search time 27 seconds  
(without alignments)  
24.562 Million cell updates/sec

Title: US-10-008-377A-2  
Perfect score: 52  
Sequence: 1 GVLMWTFSL 10

Scoring table: BIOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 110780

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A-COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5S-COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A-COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B-COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/6CTUS-COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfill1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match Length	DB	ID	Description
1	59.6	9	4	US-09-462-453-22	Sequence 22, Appl
2	53.8	9	4	US-09-667-857-438	Sequence 438, App
3	53.8	10	4	US-09-667-857-421	Sequence 421, App
4	53.8	10	4	US-09-667-857-433	Sequence 433, App
5	51.9	8	4	US-09-786-480B-42	Sequence 42, Appl
6	50.0	8	3	US-09-082-279B-148	Sequence 148B, Ap
7	50.0	8	3	US-09-315-304B-1642	Sequence 1642, Ap
8	50.0	8	4	US-09-834-784-1488	Sequence 1488, Ap
9	50.0	8	4	US-09-350-641C-1643	Sequence 1643, Ap
10	48.1	7	3	US-08-335-865-18	Sequence 18, Appl
11	46.2	6	4	US-09-121-211-12	Sequence 12, Appl
12	46.2	8	2	US-08-357-642A-4	Sequence 4, Appl
13	46.2	8	2	US-08-357-642A-14	Sequence 14, Appl
14	46.2	8	2	US-08-460-626-4	Sequence 4, Appl
15	46.2	8	2	US-08-460-626-14	Sequence 14, Appl
16	46.2	8	2	US-08-460-626-18	Sequence 18, Appl
17	46.2	8	3	US-08-836-325-4	Sequence 4, Appl
18	46.2	8	4	US-09-457-571-4	Sequence 4, Appl
19	46.2	9	3	US-08-159-339A-648	Sequence 648, App
20	46.2	9	4	US-09-682-325-22	Sequence 22, Appl
21	46.2	9	4	US-09-682-325-24	Sequence 24, Appl
22	46.2	10	2	US-08-751-300-28	Sequence 28, Appl
23	46.2	10	3	US-08-751-300-29	Sequence 29, Appl
24	46.2	10	3	US-08-159-339A-727	Sequence 727, App
25	46.2	10	4	US-08-728-742A-61	Sequence 61, Appl
26	46.2	10	4	US-09-043-813-28	Sequence 28, Appl
27	46.2	10	4	US-09-043-813-29	Sequence 29, Appl

28	23	44.2	6	1	US-07-802-667-27	Sequence 27, Appl
29	23	44.2	6	1	US-07-802-667-28	Sequence 28, Appl
30	23	44.2	6	4	US-09-282-029A-32	Sequence 32, Appl
31	23	44.2	6	4	US-09-282-029A-108	Sequence 108, Appl
32	23	44.2	6	4	US-09-185-908-32	Sequence 32, Appl
33	23	44.2	6	4	US-09-185-908-108	Sequence 108, Appl
34	23	44.2	7	4	US-09-282-029A-33	Sequence 33, Appl
35	23	44.2	7	4	US-09-282-029A-109	Sequence 109, Appl
36	23	44.2	7	4	US-09-185-908-33	Sequence 33, Appl
37	23	44.2	7	4	US-09-185-908-109	Sequence 109, Appl
38	23	44.2	8	2	US-08-641-314C-7	Sequence 7, Appl
39	23	44.2	8	3	US-09-082-279B-1433	Sequence 1433, Ap
40	23	44.2	8	3	US-09-082-279B-1474	Sequence 1474, Ap
41	23	44.2	8	3	US-09-082-279B-1484	Sequence 1484, Ap
42	23	44.2	8	3	US-09-082-279B-1485	Sequence 1485, Ap
43	23	44.2	8	3	US-09-082-279B-1487	Sequence 1487, Ap
44	23	44.2	8	3	US-09-315-304B-1433	Sequence 1433, Ap
45	23	44.2	8	3	US-09-315-304B-1562	Sequence 1562, Ap

# ALIGNMENTS

RESULT 1  
US-09-462-453-22  
Sequence 22, Application US/09462453  
Patent No. 6723695  
GENERAL INFORMATION:  
APPLICANT: BURETTS, SCOTT RENTON  
APPLICANT: KHANNA, RAJIV  
APPLICANT: SHERITT, MARTINA ALISON  
TITLE OF INVENTION: CTU EPITOPES FROM EBV  
FILE REFERENCE: FBRC:010  
CURRENT APPLICATION NUMBER: US/09/462,453  
CURRENT FILING DATE: 2000-05-30  
NUMBER OF SEQ ID NOS: 33  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 22  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Epstein Barr Virus  
US-09-462-453-22

Query Match  
Best Local Similarity 59.6%; Score 31; DB 4; Length 9;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db  
4 VILMEV 9

RESULT 2  
US-09-667-857-438  
Sequence 438, Application US/09667857  
Patent No. 669664  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: King, Gordon E.  
APPLICANT: Algate, Paul A.  
APPLICANT: Fling, Steven P.  
APPLICANT: Retter, Marc W.  
APPLICANT: Ranger, Gary Richard  
APPLICANT: Reed, Steven G.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Cartier, Darrick  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.462C5  
CURRENT APPLICATION NUMBER: US/09/667,857  
CURRENT FILING DATE: 2000-09-20  
NUMBER OF SEQ ID NOS: 455  
SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 438  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-667-857-438

Query Match  
Best Local Similarity 53.8%; Score 28; DB 4; Length 9;  
Best Local Similarity 44.4%; Pred. No. 3.8e+05;  
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 VLMEIFSL 10  
: |||||:  
Db 1 ILFWSIISI 9

RESULT 3  
US-09-667-857-421  
Sequence 421, Application US/09667857  
Patent No. 6699664  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: King, Gordon E.  
APPLICANT: Algate, Paul A.  
APPLICANT: Fling, Steven P.  
APPLICANT: Retter, Marc W.  
APPLICANT: Ranger, Gary Richard  
APPLICANT: Reed, Steven G.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darlick  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER  
FILE REFERENCE: 210121.462CS  
CURRENT APPLICATION NUMBER: US/09/667,857  
CURRENT FILING DATE: 2000-09-20  
NUMBER OF SEQ ID NOS: 455  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 421  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-667-857-421

Query Match  
Best Local Similarity 53.8%; Score 28; DB 4; Length 10;  
Best Local Similarity 44.4%; Pred. No. 63;  
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 VLMEIFSL 10  
: |||||:  
Db 1 ILFWSIISI 9

RESULT 4  
US-09-667-857-433  
Sequence 433, Application US/09667857  
Patent No. 6699664  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: King, Gordon E.  
APPLICANT: Algate, Paul A.  
APPLICANT: Fling, Steven P.  
APPLICANT: Retter, Marc W.  
APPLICANT: Ranger, Gary Richard  
APPLICANT: Reed, Steven G.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darlick  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER  
FILE REFERENCE: 210121.462CS  
CURRENT APPLICATION NUMBER: US/09/667,857  
CURRENT FILING DATE: 2000-09-20  
NUMBER OF SEQ ID NOS: 455  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 433

LENGTH: 10  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-667-857-433

Query Match  
Best Local Similarity 53.8%; Score 28; DB 4; Length 10;  
Best Local Similarity 44.4%; Pred. No. 63;  
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 VLMEIFSL 10  
: |||||:  
Db 2 ILFWSIISI 10

RESULT 5  
US-09-786-480B-42  
Sequence 42, Application US/09786480B  
Patent No. 6730825  
GENERAL INFORMATION:  
APPLICANT: Goldsbrough, Andrew  
APPLICANT: Colliver, Steve  
TITLE OF INVENTION: Isoforms of Starch Branching Enzyme II (SBE-IIA and SBE-IIB) From  
FILE REFERENCE: 11951.0005.PCUS00 MS18:005  
CURRENT APPLICATION NUMBER: US/09/786,480B  
CURRENT FILING DATE: 2002-05-25  
PRIOR APPLICATION NUMBER: PCT/GB99/03011  
PRIOR FILING DATE: 1999-09-09  
PRIOR APPLICATION NUMBER: EP 98307337.0  
PRIOR FILING DATE: 1998-09-10  
NUMBER OF SEQ ID NOS: 55  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 42  
LENGTH: 8  
TYPE: PRT  
ORGANISM: Triticum aestivum  
US-09-786-480B-42

Query Match  
Best Local Similarity 51.9%; Score 27; DB 4; Length 8;  
Best Local Similarity 80.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 LWEIF 8  
: |||||:  
Db 2 VMEIF 6

RESULT 6  
US-09-082-279B-1488  
Sequence 1488, Application US/09082279B  
Patent No. 6258782  
GENERAL INFORMATION:  
APPLICANT: Barney, Shawn  
APPLICANT: Guthe, Kelly  
APPLICANT: Wernicka, Gene  
APPLICANT: Anwer, Mohamed  
APPLICANT: Lambert, Dennis  
TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED  
TITLE OF INVENTION: PHARMACOKINETIC PROPERTIES  
FILE REFERENCE: 7872-043  
CURRENT APPLICATION NUMBER: US/09/082,279B  
CURRENT FILING DATE: 1998-05-20  
NUMBER OF SEQ ID NOS: 1515  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1488  
LENGTH: 8  
TYPE: PRT  
ORGANISM: HIV-2  
US-09-082-279B-1488

Query Match  
Best Local Similarity 50.0%; Score 26; DB 3; Length 8;  
Best Local Similarity 60.0%; Pred. No. 3.8e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 WEIFS 9  
|:|:|  
Db 1 WDVFS 5

RESULT 7  
US-09-315-304B-1642  
Sequence 1642, Application US/09315304B

GENERAL INFORMATION:  
APPLICANT: Barney, S.  
APPLICANT: Guthrie, K.  
APPLICANT: Merutka, G.  
APPLICANT: Anwer, M.  
APPLICANT: Lambert, D.  
TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC  
FILE REFERENCE: 7872-052  
CURRENT APPLICATION NUMBER: US/09/315,304B  
PRIOR FILING DATE: 1999-05-20  
PRIOR APPLICATION NUMBER: 09/082,279  
NUMBER OF SEQ ID NOS: 1667  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1642  
LENGTH: 8  
TYPE: PRT  
ORGANISM: HIV-2  
US-09-315-304B-1642

Query Match 50.0%; Score 26; DB 3; Length 8;  
Best Local Similarity 60.0%; Pred. No. 3.8e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 WEIFS 9  
|:|:|  
Db 1 WDVFS 5

RESULT 8  
US-09-834-784-1488  
Sequence 1488, Application US/09834784

GENERAL INFORMATION:  
APPLICANT: Barney, Shawn  
APPLICANT: Guthrie, Kelly  
APPLICANT: Merutka, Gene  
APPLICANT: Anwer, Mohamed  
APPLICANT: Lambert, Dennis  
TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED  
FILE REFERENCE: 7872-043  
CURRENT APPLICATION NUMBER: US/09/834,784  
PRIOR FILING DATE: 2001-04-13  
PRIOR APPLICATION NUMBER: 09/082,279  
NUMBER OF SEQ ID NOS: 1515  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1488  
LENGTH: 8  
TYPE: PRT  
ORGANISM: HIV-2  
US-09-834-784-1488

Query Match 50.0%; Score 26; DB 4; Length 8;  
Best Local Similarity 60.0%; Pred. No. 3.8e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 WEIFS 9  
|:|:|  
Db 1 WDVFS 5

RESULT 9  
US-09-350-641C-1643  
Sequence 1643, Application US/09350641C

GENERAL INFORMATION:  
APPLICANT: Barney, S.  
APPLICANT: Guthrie, K.  
APPLICANT: Merutka, G.  
APPLICANT: Anwer, M.  
APPLICANT: Lambert, D.  
TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC  
FILE REFERENCE: 7872-067  
CURRENT APPLICATION NUMBER: US/09/350,641C  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: 09/315,304  
PRIOR FILING DATE: 1999-05-20  
PRIOR APPLICATION NUMBER: 09/082,279  
NUMBER OF SEQ ID NOS: 1757  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1643  
LENGTH: 8  
TYPE: PRT  
ORGANISM: HIV-2  
US-09-350-641C-1643

Query Match 50.0%; Score 26; DB 4; Length 8;  
Best Local Similarity 60.0%; Pred. No. 3.8e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 WEIFS 9  
|:|:|  
Db 1 WDVFS 5

RESULT 10  
US-08-335-865T-18  
Sequence 18, Application US/08335865T

GENERAL INFORMATION:  
APPLICANT: Stacker, Steven A.; Hovens, Christopher M.,  
APPLICANT: Wilks, Andrew F.  
TITLE OF INVENTION: RECEPTOR-TYPE TYROSINE KINASE-LIKE MOLECULES  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Fulbright & Jaworski L.L.P.  
STREET: 666 Fifth Ave  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 720 KB storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: ASCII/Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/335,865T  
FILING DATE: 19-January-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU93/00210  
FILING DATE: 10-May-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PL2358  
FILING DATE: 11-May-1992  
ATC/NEY/AGENT INFORMATION:  
NAME: Hanson, No. 6107472man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD-5277  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 318-3100

TELEFAX: (212) 752-5958  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-357-8657-18

Query Match 48.1%; Score 25; DB 3; Length 7;  
Best Local Similarity 66.7%; Pred. No. 3.8e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLWEL 7  
DB 1 VTLWEL 6

RESULT 11  
US-09-121-211-12  
Sequence 12, Application US/09121211  
Patent No. 675052  
GENERAL INFORMATION:  
APPLICANT: Shinohara, Toshimichi  
APPLICANT: Shingh, Dharendra P.  
APPLICANT: Chylack, Leo T.  
TITLE OF INVENTION: Lens Epithelial Cell Derived Growth  
FILE REFERENCE: B0801/7116  
CURRENT APPLICATION NUMBER: US/09/121,211  
EARLIER FILING DATE: 1998-07-23  
EARLIER APPLICATION NUMBER: U.S. 60/053,549  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 12  
LENGTH: 6  
TYPE: PRT  
ORGANISM: Homo Sapiens  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (1)...(6)  
US-09-121-211-12

Query Match 46.2%; Score 24; DB 4; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LWEL 7  
DB 3 LWEL 6

RESULT 12  
US-08-357-642A-4  
Sequence 4, Application US/08357642A  
Patent No. 5837524  
GENERAL INFORMATION:  
APPLICANT: Sina Lev  
APPLICANT: Joseph Schlessinger  
TITLE OF INVENTION: PYK2 RELATED PRODUCTS  
NUMBER OF INVENTIONS: AND METHODS  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

MEDIUM TYPE: storage  
COMPUTER: IBM Compatib1e  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/357,642A  
FILING DATE: December 15, 1994  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 209/070  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-357-642A-4

Query Match 46.2%; Score 24; DB 2; Length 8;  
Best Local Similarity 80.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVLW 5  
DB 4 GVLW 8

RESULT 13  
US-08-357-642A-14  
Sequence 14, Application US/08357642A  
Patent No. 5837524  
GENERAL INFORMATION:  
APPLICANT: Sina Lev  
APPLICANT: Joseph Schlessinger  
TITLE OF INVENTION: PYK2 RELATED PRODUCTS  
NUMBER OF INVENTIONS: AND METHODS  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatib1e  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/357,642A  
FILING DATE: December 15, 1994  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 209/070  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-357-642a-14

Query Match 46.2%; Score 24; DB 2; Length 8;  
Best Local Similarity 80.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVLM 5  
Db 4 GVLM 8

RESULT 14  
US-08-460-626-4  
Sequence 4, Application US/08460626  
Patent No. 5837815  
GENERAL INFORMATION:  
APPLICANT: SIMA LEV  
TITLE OF INVENTION: PYK-2 RELATED PRODUCTS AND  
METHODS  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,626  
FILING DATE: June 2, 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/357,642  
FILING DATE: December 15, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 211/121  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-460-626-4

Query Match 46.2%; Score 24; DB 2; Length 8;  
Best Local Similarity 80.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVLM 5  
Db 4 GVLM 8

RESULT 15  
US-08-460-626-14  
Sequence 14, Application US/08460626  
Patent No. 5837815  
GENERAL INFORMATION:  
APPLICANT: SIMA LEV  
TITLE OF INVENTION: PYK-2 RELATED PRODUCTS AND  
METHODS  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,626  
FILING DATE: June 2, 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/357,642  
FILING DATE: December 15, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 211/121  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-460-626-14

Query Match 46.2%; Score 24; DB 2; Length 8;  
Best Local Similarity 80.0%; Pred. No. 3.8e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVLM 5  
Db 4 GVLM 8

Search completed: November 30, 2004, 08:15:58  
Job time : 27 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: November 30, 2004, 08:13:53 ; Search time 44 Seconds  
(without alignments)  
80.612 Million cell updates/sec

Title: US-10-008-377a-2

Perfect score: 52

Sequence: 1 GVILMEIFSL 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1575965 seqs, 354694765 residues

Total number of hits satisfying chosen parameters: 177356

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database:

Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubppa/PC7\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubppa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubppa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubppa/PC7US\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubppa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubppa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubppa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubppa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubppa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubppa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	10	14	US-10-008-377-2
2	46	88.5	9	14	US-10-008-377-5
3	32	61.5	10	10	US-09-572-404B-1154
4	31	59.6	9	14	US-10-356-829-22
5	29	55.8	9	15	US-10-149-138-360
6	29	55.8	9	16	US-10-149-138-360
7	29	55.8	10	15	US-10-149-138-92
8	29	55.8	10	15	US-10-149-138-1177
9	29	55.8	10	15	US-10-149-138-2106
10	29	55.8	10	15	US-10-149-138-2707
11	29	55.8	10	16	US-10-149-138-92
12	29	55.8	10	16	US-10-149-138-1177
13	29	55.8	10	16	US-10-149-138-2106

14	29	55.8	10	16	US-10-149-138-2707	Sequence 2707, Ap
15	28	53.8	8	15	US-10-149-138-359	Sequence 359, App
16	28	53.8	8	15	US-10-149-138-1176	Sequence 1176, Ap
17	28	53.8	8	15	US-10-149-138-2105	Sequence 2105, App
18	28	53.8	8	16	US-10-149-138-359	Sequence 359, App
19	28	53.8	8	16	US-10-149-138-1176	Sequence 1176, Ap
20	28	53.8	8	16	US-10-149-138-2105	Sequence 2105, Ap
21	28	53.8	9	9	US-09-884-441-438	Sequence 438, App
22	28	53.8	9	10	US-09-907-969-438	Sequence 438, App
23	28	53.8	9	10	US-09-827-271-438	Sequence 438, App
24	28	53.8	9	14	US-10-198-053-438	Sequence 438, App
25	28	53.8	9	15	US-10-245-871-606	Sequence 606, App
26	28	53.8	9	15	US-10-149-138-1375	Sequence 1375, App
27	28	53.8	9	15	US-10-149-138-3703	Sequence 3703, App
28	28	53.8	9	15	US-10-253-286-606	Sequence 606, App
29	28	53.8	9	16	US-10-149-138-1375	Sequence 1375, App
30	28	53.8	9	16	US-10-149-138-3703	Sequence 3703, App
31	28	53.8	10	9	US-09-884-441-421	Sequence 421, App
32	28	53.8	10	9	US-09-884-441-433	Sequence 433, App
33	28	53.8	10	10	US-09-907-969-421	Sequence 421, App
34	28	53.8	10	10	US-09-907-969-433	Sequence 433, App
35	28	53.8	10	10	US-09-827-271-421	Sequence 421, App
36	28	53.8	10	10	US-09-827-271-433	Sequence 433, App
37	28	53.8	10	14	US-10-198-053-421	Sequence 421, App
38	28	53.8	10	14	US-10-198-053-433	Sequence 433, App
39	28	53.8	10	15	US-10-149-138-1376	Sequence 1376, App
40	28	53.8	10	15	US-10-149-138-3000	Sequence 3000, App
41	28	53.8	10	16	US-10-149-138-1376	Sequence 1376, App
42	28	53.8	10	16	US-10-149-138-2000	Sequence 2000, App
43	27	51.9	8	17	US-10-818-770-42	Sequence 42, App1
44	27	51.9	10	14	US-10-190-082-596	Sequence 596, App
45	26	50.0	8	14	US-10-351-641-1643	Sequence 1643, App

## ALIGNMENTS

RESULT 1  
US-10-008-377-2  
; Sequence 2, Application US/10008377  
; Publication No. US20030157101A1  
; GENERAL INFORMATION:  
; APPLICANT: Gambacorti-Passerini, Carlo  
; TITLE OF INVENTION: Immunogenic ALK Peptides  
; FILE REFERENCE: 045922/241203  
; CURRENT APPLICATION NUMBER: US/10/008,377  
; CURRENT FILING DATE: 2001-11-15  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-008-377-2

Query Match 100.0%; Score 52; DB 14; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.031; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0;

QY 1 GVILMEIFSL 10  
DB 1 GVILMEIFSL 10

RESULT 2  
US-10-008-377-5  
; Sequence 5, Application US/10008377  
; Publication No. US20030157101A1  
; GENERAL INFORMATION:  
; APPLICANT: Gambacorti-Passerini, Carlo  
; APPLICANT: Passoni, Lorena  
; TITLE OF INVENTION: Immunogenic ALK Peptides

FILE REFERENCE: 045922/241203  
CURRENT APPLICATION NUMBER: US/10/008,377  
CURRENT FILING DATE: 2001-11-15  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-008-377-5

Query Match 88.5%; Score 46; DB 14; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLWEIFSL 10  
DB 1 VLWEIFSL 9

RESULT 3  
US-09-572-404B-1154  
Sequence 1154, Application US/09572404B  
Publication No. US20030078374A1  
GENERAL INFORMATION:  
APPLICANT: Proteom Ltd  
TITLE OF INVENTION: Complementary peptide ligands from the human genome  
FILE REFERENCE: Human patent  
CURRENT APPLICATION NUMBER: US/09/572,404B  
CURRENT FILING DATE: 2000-05-17  
NUMBER OF SEQ ID NOS: 4203  
SOFTWARE: Protpatent version 1.0  
SEQ ID NO 1154  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Homo Sapiens  
FEATURE:  
OTHER INFORMATION: sequence located in Unknown at 860-869 and may interact with sequ  
US-09-572-404B-1154

Query Match 61.5%; Score 32; DB 10; Length 10;  
Best Local Similarity 83.3%; Pred. No. 77;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 WEIFSL 10  
DB 1 WEIFSL 6

RESULT 4  
US-10-356-829-22  
Sequence 22, Application US/10356829  
Publication No. US20030155582A1  
GENERAL INFORMATION:  
APPLICANT: BURROWS, SCOTT RENTON  
APPLICANT: KHANN, RAJIV  
APPLICANT: SHERRITT, MARTINA ALISON  
TITLE OF INVENTION: CTL EPITOPES FROM EBV  
FILE REFERENCE: PRC:010UCSD1  
CURRENT APPLICATION NUMBER: US/10/356,829  
CURRENT FILING DATE: 2003-02-03  
PRIOR APPLICATION NUMBER: 09/462,453  
PRIOR FILING DATE: 2000-01-06  
NUMBER OF SEQ ID NOS: 33  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 22  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Epstein Barr Virus  
US-10-356-829-22

Query Match 59.6%; Score 31; DB 14; Length 9;

Best Local Similarity 83.3%; Pred. No. 1.4e+06;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLWEI 7  
DB 4 VLWEV 9

RESULT 5  
US-10-149-138-360  
Sequence 360, Application US/10149138  
Publication No. US20040018971A1  
GENERAL INFORMATION:  
APPLICANT: Fikes, John  
APPLICANT: Sette, Alessandro  
APPLICANT: Sidney, John  
APPLICANT: Southwood, Scott  
APPLICANT: Chesnut, Robert  
APPLICANT: Celis, Esben  
APPLICANT: Keogh, Elissa  
TITLE OF INVENTION: Inducing Cellular Immune Responses to  
FILE REFERENCE: 2060.0140001  
CURRENT APPLICATION NUMBER: US/10/149,138  
CURRENT FILING DATE: 2002-06-10  
PRIOR APPLICATION NUMBER: PCT/US00/33591  
PRIOR FILING DATE: 2000-12-11  
PRIOR APPLICATION NUMBER: US 09/458,299  
PRIOR FILING DATE: 1999-12-11  
NUMBER OF SEQ ID NOS: 4641  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 360  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Artificially Synthesized Peptide  
US-10-149-138-360

Query Match 55.8%; Score 29; DB 15; Length 9;  
Best Local Similarity 44.4%; Pred. No. 1.4e+06;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVLWEIFS 9  
DB 1 GVLWEIWT 9

RESULT 6  
US-10-149-138-360  
Sequence 360, Application US/10149138  
Publication No. US20040121946A9  
GENERAL INFORMATION:  
APPLICANT: Fikes, John  
APPLICANT: Sette, Alessandro  
APPLICANT: Sidney, John  
APPLICANT: Southwood, Scott  
APPLICANT: Chesnut, Robert  
APPLICANT: Celis, Esben  
APPLICANT: Keogh, Elissa  
TITLE OF INVENTION: Inducing Cellular Immune Responses to  
FILE REFERENCE: 2060.0140001  
CURRENT APPLICATION NUMBER: US/10/149,138  
CURRENT FILING DATE: 2002-06-10  
PRIOR APPLICATION NUMBER: PCT/US00/33591  
PRIOR FILING DATE: 2000-12-11  
PRIOR APPLICATION NUMBER: US 09/458,299  
PRIOR FILING DATE: 1999-12-11  
NUMBER OF SEQ ID NOS: 4641  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 360  
LENGTH: 9

TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Artificially Synthesized Peptide  
US-10-149-138-360

Query Match 55.8%; Score 29; DB 16; Length 9;  
Best Local Similarity 44.4%; Pred. No. 1.4e+06;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GYLMWEIFS 9  
DB 1 GYTWELMT 9

RESULT 7  
US-10-149-138-92  
Sequence 92, Application US/10149138  
Publication No. US20040018971A1  
GENERAL INFORMATION:  
APPLICANT: Fikes, John  
APPLICANT: Sette, Alessandro  
APPLICANT: Sidney, John  
APPLICANT: Southwood, Scott  
APPLICANT: Chesnut, Robert  
APPLICANT: Celis, Esteban  
APPLICANT: Keogh, Elissa  
TITLE OF INVENTION: Inducing Cellular Immune Responses to  
TITLE OF INVENTION: HER2/neu Using Peptide and Nucleic Acid Compositions  
FILE REFERENCE: 2060.0140001  
CURRENT APPLICATION NUMBER: US/10/149,138  
CURRENT FILING DATE: 2002-06-10  
PRIOR APPLICATION NUMBER: PCT/US00/33591  
PRIOR FILING DATE: 2000-12-11  
PRIOR APPLICATION NUMBER: US 09/458,299  
PRIOR FILING DATE: 1999-12-11  
NUMBER OF SEQ ID NOS: 4641  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 92  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Artificially Synthesized Peptide  
US-10-149-138-92

Query Match 55.8%; Score 29; DB 15; Length 10;  
Best Local Similarity 44.4%; Pred. No. 2.5e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GYLMWEIFS 9  
DB 1 GYTWELMT 9

RESULT 8  
US-10-149-138-1177  
Sequence 1177, Application US/10149138  
Publication No. US20040018971A1  
GENERAL INFORMATION:  
APPLICANT: Fikes, John  
APPLICANT: Sette, Alessandro  
APPLICANT: Sidney, John  
APPLICANT: Southwood, Scott  
APPLICANT: Chesnut, Robert  
APPLICANT: Celis, Esteban  
APPLICANT: Keogh, Elissa  
TITLE OF INVENTION: Inducing Cellular Immune Responses to  
TITLE OF INVENTION: HER2/neu Using Peptide and Nucleic Acid Compositions  
FILE REFERENCE: 2060.0140001  
CURRENT APPLICATION NUMBER: US/10/149,138  
CURRENT FILING DATE: 2002-06-10  
PRIOR APPLICATION NUMBER: PCT/US00/33591

PRIOR FILING DATE: 2000-12-11  
PRIOR APPLICATION NUMBER: US 09/458,299  
PRIOR FILING DATE: 1999-12-11  
NUMBER OF SEQ ID NOS: 4641  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1177  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Artificially Synthesized Peptide  
US-10-149-138-1177

Query Match 55.8%; Score 29; DB 15; Length 10;  
Best Local Similarity 44.4%; Pred. No. 2.5e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GYLMWEIFS 9  
DB 1 GYTWELMT 9

RESULT 9  
US-10-149-138-2106  
Sequence 2106, Application US/10149138  
Publication No. US20040018971A1  
GENERAL INFORMATION:  
APPLICANT: Fikes, John  
APPLICANT: Sette, Alessandro  
APPLICANT: Sidney, John  
APPLICANT: Southwood, Scott  
APPLICANT: Chesnut, Robert  
APPLICANT: Celis, Esteban  
APPLICANT: Keogh, Elissa  
TITLE OF INVENTION: Inducing Cellular Immune Responses to  
TITLE OF INVENTION: HER2/neu Using Peptide and Nucleic Acid Compositions  
FILE REFERENCE: 2060.0140001  
CURRENT APPLICATION NUMBER: US/10/149,138  
CURRENT FILING DATE: 2002-06-10  
PRIOR APPLICATION NUMBER: PCT/US00/33591  
PRIOR FILING DATE: 2000-12-11  
PRIOR APPLICATION NUMBER: US 09/458,299  
PRIOR FILING DATE: 1999-12-11  
NUMBER OF SEQ ID NOS: 4641  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2106  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Artificially Synthesized Peptide  
US-10-149-138-2106

Query Match 55.8%; Score 29; DB 15; Length 10;  
Best Local Similarity 44.4%; Pred. No. 2.5e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GYLMWEIFS 9  
DB 1 GYTWELMT 9

RESULT 10  
US-10-149-138-2707  
Sequence 2707, Application US/10149138  
Publication No. US20040018971A1  
GENERAL INFORMATION:  
APPLICANT: Fikes, John  
APPLICANT: Sette, Alessandro  
APPLICANT: Sidney, John  
APPLICANT: Southwood, Scott  
APPLICANT: Chesnut, Robert  
APPLICANT: Celis, Esteban

APPLICANT: Keogh, Elissa  
TITLE OF INVENTION: Inducing Cellular Immune Responses to  
FILE REFERENCE: 2060.0140001  
CURRENT FILING DATE: 2002-06-10  
PRIOR APPLICATION NUMBER: PCT/US00/33591  
PRIOR FILING DATE: 2000-12-11  
PRIOR APPLICATION NUMBER: US 09/458,299  
NUMBER OF SEQ ID NOS: 4641  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2707  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Artificially Synthesized Peptide  
US-10-149-138-2707

Query Match 55.8%; Score 29; DB 15; Length 10;  
Best Local Similarity 44.4%; Pred. No. 2.5e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 GVLMWEIFS 9  
DB 1 GVTWELMT 9

RESULT 11  
US-10-149-138-92  
Sequence 92, Application US/10149138  
Publication No. US20040121946A9  
GENERAL INFORMATION:  
APPLICANT: Fikes, John  
APPLICANT: Sette, Alessandro  
APPLICANT: Sidney, John  
APPLICANT: Southwood, Scott  
APPLICANT: Chesnut, Robert  
APPLICANT: Celis, Esteban  
APPLICANT: Keogh, Elissa  
TITLE OF INVENTION: Inducing Cellular Immune Responses to  
FILE REFERENCE: 2060.0140001  
CURRENT FILING DATE: 2002-06-10  
PRIOR APPLICATION NUMBER: PCT/US00/33591  
PRIOR FILING DATE: 2000-12-11  
PRIOR APPLICATION NUMBER: US 09/458,299  
NUMBER OF SEQ ID NOS: 4641  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 92  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Artificially Synthesized Peptide  
US-10-149-138-92

Query Match 55.8%; Score 29; DB 16; Length 10;  
Best Local Similarity 44.4%; Pred. No. 2.5e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 GVLMWEIFS 9  
DB 1 GVTWELMT 9

RESULT 12  
US-10-149-138-1177  
Sequence 1177, Application US/10149138  
Publication No. US20040121946A9

GENERAL INFORMATION:  
APPLICANT: Fikes, John  
APPLICANT: Sette, Alessandro  
APPLICANT: Sidney, John  
APPLICANT: Southwood, Scott  
APPLICANT: Chesnut, Robert  
APPLICANT: Celis, Esteban  
APPLICANT: Keogh, Elissa  
TITLE OF INVENTION: Inducing Cellular Immune Responses to  
FILE REFERENCE: 2060.0140001  
CURRENT FILING DATE: 2002-06-10  
PRIOR APPLICATION NUMBER: PCT/US00/33591  
PRIOR FILING DATE: 2000-12-11  
PRIOR APPLICATION NUMBER: US 09/458,299  
NUMBER OF SEQ ID NOS: 4641  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1177  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Artificially Synthesized Peptide  
US-10-149-138-1177

Query Match 55.8%; Score 29; DB 16; Length 10;  
Best Local Similarity 44.4%; Pred. No. 2.5e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 GVLMWEIFS 9  
DB 1 GVTWELMT 9

RESULT 13  
US-10-149-138-2106  
Sequence 2106, Application US/10149138  
Publication No. US20040121946A9  
GENERAL INFORMATION:  
APPLICANT: Fikes, John  
APPLICANT: Sette, Alessandro  
APPLICANT: Sidney, John  
APPLICANT: Southwood, Scott  
APPLICANT: Chesnut, Robert  
APPLICANT: Celis, Esteban  
APPLICANT: Keogh, Elissa  
TITLE OF INVENTION: Inducing Cellular Immune Responses to  
FILE REFERENCE: 2060.0140001  
CURRENT FILING DATE: 2002-06-10  
PRIOR APPLICATION NUMBER: PCT/US00/33591  
PRIOR FILING DATE: 2000-12-11  
PRIOR APPLICATION NUMBER: US 09/458,299  
NUMBER OF SEQ ID NOS: 4641  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2106  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Artificially Synthesized Peptide  
US-10-149-138-2106

Query Match 55.8%; Score 29; DB 16; Length 10;  
Best Local Similarity 44.4%; Pred. No. 2.5e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 GVLMWEIFS 9  
DB 1 GVTWELMT 9

Db 1 GVTWELMT 9

RESULT 14

US-10-149-138-2707

; Sequence 2707, Application US/10149138

; Publication No. US20040121946A9

; GENERAL INFORMATION:

; APPLICANT: Fikes, John

; APPLICANT: Sette, Alessandro

; APPLICANT: Sidney, John

; APPLICANT: Southwood, Scott

; APPLICANT: Chesnut, Robert

; APPLICANT: Celis, Esteban

; APPLICANT: Keogh, Elissa

; TITLE OF INVENTION: Inducing Cellular Immune Responses to

; FILE REFERENCE: 2060.0140001

; CURRENT APPLICATION NUMBER: US/10/149,138

; CURRENT FILING DATE: 2002-06-10

; PRIOR APPLICATION NUMBER: PCT/US00/33591

; PRIOR FILING DATE: 2000-12-11

; PRIOR APPLICATION NUMBER: US 09/458,299

; PRIOR FILING DATE: 1999-12-11

; NUMBER OF SEQ ID NOS: 4641

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2707

; LENGTH: 10

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Artificially Synthesized Peptide

US-10-149-138-2707

Query Match 55.8%; Score 29; DB 16; Length 10;

Best Local Similarity 44.4%; Pred. No. 2.5e+02;

Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVTWELMT 9

Db 1 GVTWELMT 9

RESULT 15

US-10-149-138-359

; Sequence 359, Application US/10149138

; Publication No. US20040018971A1

; GENERAL INFORMATION:

; APPLICANT: Fikes, John

; APPLICANT: Sette, Alessandro

; APPLICANT: Sidney, John

; APPLICANT: Southwood, Scott

; APPLICANT: Chesnut, Robert

; APPLICANT: Celis, Esteban

; APPLICANT: Keogh, Elissa

; TITLE OF INVENTION: Inducing Cellular Immune Responses to

; FILE REFERENCE: 2060.0140001

; CURRENT APPLICATION NUMBER: US/10/149,138

; CURRENT FILING DATE: 2002-06-10

; PRIOR APPLICATION NUMBER: PCT/US00/33591

; PRIOR FILING DATE: 2000-12-11

; PRIOR APPLICATION NUMBER: US 09/458,299

; PRIOR FILING DATE: 1999-12-11

; NUMBER OF SEQ ID NOS: 4641

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 359

; LENGTH: 8

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Artificially Synthesized Peptide

US-10-149-138-359

Query Match 53.8%; Score 28; DB 15; Length 8;

Best Local Similarity 57.1%; Pred. No. 1.4e+06;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVTWELMT 7

Db 1 GVTWELMT 7

Search completed: November 30, 2004, 08:31:00  
Job time : 44 secs

This Page Blank (uspto)

Tue Nov.30 08:49:17 2004

us-10-008-377a-1.closed.rpx

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:41:32; Search time 11.6667 Seconds  
(without alignments)  
82.472 Million cell updates/sec

Title: US-10-008-377A-1

Perfect score: 47

Sequence: 1 SLAMDLHV 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 28316 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 1102

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: PIR1:  
2: PIR2:  
3: PIR3:  
4: PIR4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	23	48.9	10 2	E86128
2	19	40.4	10 2	D33098
3	17	36.2	10 2	PIR1345
4	16	34.0	6 2	S78764
5	16	34.0	10 2	C26997
6	16	34.0	10 2	A43977
7	15	31.9	7 2	S20446
8	15	31.9	10 2	B46453
9	15	31.9	10 2	A56633
10	15	31.9	10 2	A32543
11	14	29.8	7 2	UN0859
12	14	29.8	7 2	PC0663
13	14	29.8	7 2	A28340
14	14	29.8	10 2	B43590
15	14	29.8	10 2	J00943
16	14	29.8	10 2	A43590
17	14	29.8	10 2	S27873
18	13	27.7	7 2	S55548
19	13	27.7	7 2	A28709
20	13	27.7	7 2	ECMUCR
21	13	27.7	8 2	S66296
22	13	27.7	9 2	B45796
23	13	27.7	9 2	S78426
24	13	27.7	10 2	T40251
25	13	27.7	10 2	T40332
26	13	27.7	10 2	A30823
27	12	25.5	6 2	T10954
28	12	25.5	5 2	A46474
29	12	25.5	7 2	S25266

30	12	25.5	7 2	PT0520
31	12	25.5	7 2	S29735
32	12	25.5	7 2	PC2370
33	12	25.5	8 2	S68802
34	12	25.5	8 2	PT0368
35	12	25.5	8 2	PT0298
36	12	25.5	9 2	E28854
37	12	25.5	9 2	S65913
38	12	25.5	9 2	A33527
39	12	25.5	9 2	S02384
40	12	25.5	9 2	A60108
41	12	25.5	9 2	D48186
42	12	25.5	9 2	S55696
43	12	25.5	9 2	JN0025
44	12	25.5	10 2	A61354
45	12	25.5	10 2	S13224

## ALIGNMENTS

### RESULT 1

hypothetical protein Z5903 (imported) - Escherichia coli (strain O157:H7, substrain EDL486128)  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: E86128  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
Miller, L.; Grobeck, E.J.; Davis, N.W.; Linn, A.; Dimalanta, E.; Potamoultis, K.; Apodaca  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A65480; NCID:21074935; PMID:11206551  
A:Accession: E86128  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-10 <STO>  
A:Cross-references: UNIPROT:Q8X4E5; GB:AE005174; NID:G12519314; PIDN:ACG59489.1; GSPDB:Q  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: Z5903

Query Match 48.9%; Score 23; DB 2; Length 10;  
Best Local Similarity 57.1%; Pred. No. 1.4e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 MLMDLHV 10  
DB 1 MIDLIQV 7

### RESULT 2

214K exantigen (version 1) - malaria parasite (Plasmodium falciparum) (fragments)  
C:Species: Plasmodium falciparum  
C:Date: 24-Aug-1990 #sequence\_revision 24-Aug-1990 #text\_change 09-Jun-2000  
C:Accession: D33098  
R:Nichols, J.H.; Hager, L.P.  
submitted to the Protein Sequence Database, May 1990  
A:Reference number: A33098  
A:Accession: D33098  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-10 <NIC>

Query Match 40.4%; Score 19; DB 2; Length 10;  
Best Local Similarity 80.0%; Pred. No. 9.1e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAMLD 6  
DB 2 LAMLD 6

## RESULT 3

PH1345

Ig heavy chain DJ region (clone C100-94) - human (fragment)

C/Species: Homo sapiens (man)

C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999

C/Accession: PH1345

R/Masserian, R.; Gallili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.

J. Exp. Med. 176, 1577-1581, 1992

A/Title: Predominance of fetal type DJH joining in young children with B precursor lymph

A/Reference number: PH1302; PMID:93094761; PMID:1460419

A/Accession: PH1345

A/Molecule type: DNA

A/Residues: 1-10 &lt;MAS&gt;

A/Note: the authors translated the stop codon for residue 4 as X

C/Keywords: heterotetramer; immunoglobulin

QY

3 AMDLL 8

DB

1 AMVXLL 6

Matches

4; Conservative

1; Mismatches

1; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

DB

1 MDLL 4

Matches

4; Conservative

1; Mismatches

1; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

DB

1 MDLL 4

Matches

4; Conservative

1; Mismatches

1; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

DB

1 MDLL 4

Matches

4; Conservative

1; Mismatches

1; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

DB

1 MDLL 4

Matches

4; Conservative

1; Mismatches

1; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

DB

1 MDLL 4

Matches

4; Conservative

1; Mismatches

1; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

DB

1 MDLL 4

Matches

4; Conservative

1; Mismatches

1; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

0; Mismatches

A/Note: sequence extracted from NCBI backbone (NCBIP:60243)  
F:1-10/Domain: signal sequence (fragment) #status predicted <SIG>

Query Match 31.9%; Score 15; DB 2; Length 10;  
Best Local Similarity 33.3%; Pred. No. 6e+03;  
Matches 2; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 DLMHV 10  
1 MQLFHL 6

RESULT 9  
A56633  
neomycinopressin - flesh fly (Sarcophaga bullata)

N/Alternate names: Neb-MS  
C/Species: Sarcophaga bullata  
C/Date: 21-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Jul-2004  
C/Accession: A56633  
R/Fonagy, A.; Schoofs, L.; Proost, P.; Van Damme, J.; Bueds, H.; De Loof, A.  
Comp. Biochem. Physiol. C 102, 239-245, 1992  
A/Title: Isolation, primary structure and synthesis of neomycinopressin, a myoinhibiting  
A/Reference number: A56633; PMID:93047886; PMID:1356537  
A/Accession: A56633

A/Molecule type: protein  
A/Residues: 1-10 <FON>  
A/Cross-references: UNIPROT:P61850  
A/Experimental source: head  
A/Note: sequence extracted from NCBI backbone (NCBIP:119072)  
C/Keywords: amidated carboxyl end; neuropeptide  
F:10/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 31.9%; Score 15; DB 2; Length 10;  
Best Local Similarity 60.0%; Pred. No. 6e+03;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 DLMHV 10  
2 DVDHV 6

RESULT 10  
A32543  
cardioexcitatory neuropeptide - desert locust  
C/Species: Schistocerca gregaria (desert locust)  
C/Date: 20-Dec-1989 #sequence\_revision 20-Dec-1989 #text\_change 09-Jul-2004  
C/Accession: A32543  
R/Robb, S.; Packman, L.C.; Evans, P.D.  
Biochem. Biophys. Res. Commun. 160, 850-856, 1989  
A/Title: Isolation, primary structure and bioactivity of Schistocerca gregaria, a FMRF-amide  
A/Reference number: A32543; PMID:89246543; PMID:2719702  
A/Accession: A32543  
A/Status: Preliminary  
A/Molecule type: protein  
A/Residues: 1-10 <ROB>  
A/Cross-references: UNIPROT:P38553  
C/Suprafamily: unassigned animal peptides  
C/Keywords: amidated carboxyl end; neuropeptide  
F:10/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 31.9%; Score 15; DB 2; Length 10;  
Best Local Similarity 60.0%; Pred. No. 6e+03;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 DLMHV 10  
2 DVDHV 6

RESULT 11  
JN0859  
peptidyl-dipeptidase A inhibitory peptide C105 - striped bonito  
C/Species: Sarda orientalis (striped bonito)

C/Date: 10-Mar-1994 #sequence\_revision 10-Mar-1994 #text\_change 07-May-1999  
C/Accession: JN0859  
R/Matsumura, N.; Fujii, M.; Takeda, Y.; Shimizu, T.  
Biochim. Biophys. Acta 1174, 1743-1744, 1993  
A/Title: Isolation and characterization of angiotensin I-converting enzyme inhibitory P  
A/Reference number: JN0859; PMID:94080036; PMID:7764272  
A/Accession: JN0859  
A/Molecule type: protein  
A/Residues: 1-7 <MAT>  
A/Experimental source: intestine  
A/Comment: The carboxyl-terminus is essential for the protein's expression of angiotensin  
C/Suprafamily: bradykinin-potentiating peptide  
C/Keywords: angiotensin-converting enzyme inhibitor

Query Match 29.8%; Score 14; DB 2; Length 7;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMD 6  
1 SVAKLE 6

RESULT 12

P00663  
membrane protein - porcine epidemic diarrhea virus (isolate Belgian CV777) (fragment)  
C/Species: porcine epidemic diarrhea virus  
C/Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 08-Oct-1999  
C/Accession: P00663  
R/Bridgen, A.; Duarte, M.; Tobler, K.; Laude, H.; Ackermann, M.  
J. Gen. Virol. 74, 1795-1804, 1993  
A/Title: Sequence determination of the nucleocapsid protein gene of the porcine epidemic  
A/Reference number: J02191; PMID:93389433; PMID:8397280  
A/Accession: P00663  
A/Molecule type: mRNA  
A/Residues: 1-7 <BR>  
A/Cross-references: GB:Z14976; NID:G311650; PIDN:CAAT8699.1; PID:9584083  
C/Comment: This virus is coronavirus related to human coronavirus 229E.  
C/Keywords: membrane protein

Query Match 29.8%; Score 14; DB 2; Length 7;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 7 LHMV 10  
3 VLHV 6

RESULT 13  
A28340  
myomodulin - California sea hare  
C/Species: Aplysia californica (California sea hare)  
C/Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 09-Jul-2004  
C/Accession: A28340  
R/Cropper, E.C.; Tenenbaum, R.; Kols, M.A.G.; Kimpfermann, I.; Weiss, K.R.  
Proc. Natl. Acad. Sci. U.S.A. 84, 5483-5486, 1987  
A/Title: Myomodulin: A bioactive neuropeptide present in an identified cholinergic buccal  
A/Reference number: A28340; PMID:87261010; PMID:3474664  
A/Accession: A28340  
A/Molecule type: protein  
A/Residues: 1-7 <CRO>  
A/Cross-references: UNIPROT:P15513

Query Match 29.8%; Score 14; DB 2; Length 7;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMCDL 7  
2 MSMLRL 7

## RESULT 14

B43590

pilin type A66 - Aeromonas hydrophila (fragment)

C:Species: Aeromonas hydrophila

C:Date: 12-Jan-1993 #sequence\_revision 12-Jan-1993 #text\_change 17-Nov-2000

C:Accession: B43590

R:Hokama, A.; Iwanaga, M.

Infect. Immun. 59, 3478-3483, 1991

A:Title: Purification and characterization of Aeromonas sobria pilin, a possible colonizer

A:Reference number: A43590; MUID:91372953; PMID:1680105

A:Accession: B43590

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-10 &lt;OK&gt;

C:Superfamily: Vibrio cholerae prepilin-like 17.3K protein type 4

Query Match

Best Local Similarity 29.8%; Score 14; DB 2; Length 10;

Matches 2; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 LAMDLHV 10

DB 1 MTLIELVIV 9

## RESULT 15

J00943

hypothetical 1.3K protein - infectious bursal disease virus (strain 52/70)

C:Species: infectious bursal disease virus

C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004

C:Accession: J00943

R:Baylis, C.D.; Spies, U.; Shaw, K.; Peters, R.W.; Papageorgiou, A.; Mueller, H.; Bourc

U. Gen. Virol. 71, 1303-1312, 1990

A:Title: A comparison of the sequences of segment A of four infectious bursal disease vi

A:Reference number: J00941; MUID:90278420; PMID:2161902

A:Accession: J00943

A:Status: translation not shown

A:Molecule type: genomic RNA

A:Residues: 1-10 &lt;BAV&gt;

A:Cross-references: UNIPROT:Q82625; GB:000869; NID:9221036; PIDN:BAA00743.1; PID:dl00120

Query Match

Best Local Similarity 29.8%; Score 14; DB 2; Length 10;

Matches 3; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 LAMDLHV 9

DB 3 LLLLRVH 10

Search completed: November 30, 2004, 08:14:25  
 Job time : 12.6667 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:40:12 ; Search time 78.3333 Seconds  
(without alignments)  
73.452 Million cell updates/sec

Title: US-10-008-377a-1  
Perfect score: 47  
Sequence: 1 SLAMDLLHV 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues  
Total number of hits satisfying chosen parameters: 2971

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: UniProt\_02:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	23	48.9	10 2 Q8X4E5	Q8X4E5 escherichia
2	20	42.6	8 2 Q37854	Q37854 bacteriophage
3	18	38.3	8 2 Q51594	Q51594 plasmid col
4	18	38.3	9 1 P842_LITRU	P84205 listeria rub
5	18	38.3	9 2 Q7RDE8	Q7RDE8 plasmodium
6	17	36.2	7 2 Q66205	Q66205 transmissib
7	16	34.0	8 2 Q9T778	Q9T778 canis famli
8	16	34.0	9 2 Q70SM2	Q70SM2 homo sapien
9	16	34.0	9 2 Q6G126	Q6G126 lactobacilli
10	16	34.0	9 2 Q9QZAV	Q9QZAV mus musculu
11	16	34.0	9 2 Q8UTD7	Q8UTD7 human immun
12	16	34.0	9 2 CAD79467	CAD79467 homo sapi
13	16	34.0	10 2 P8583	P8583 manduca sex
14	16	34.0	10 2 Q9T5Z3	Q9T5Z3 sus scrofa
15	16	34.0	10 2 Q39948	Q39948 gb virus c/
16	16	34.0	10 2 Q39958	Q39958 gb virus c/
17	16	34.0	10 2 Q90344	Q90344 gb virus c/
18	16	34.0	10 2 Q77VW7	Q77VW7 gb virus c/
19	16	34.0	10 2 Q77VW8	Q77VW8 gb virus c/
20	16	34.0	10 2 Q77ZJ1	Q77ZJ1 gb virus c/
21	16	34.0	10 2 Q77ZJ2	Q77ZJ2 gb virus c/
22	16	34.0	10 2 Q77ZJ3	Q77ZJ3 gb virus c/
23	16	34.0	10 2 Q77ZJ4	Q77ZJ4 gb virus c/
24	16	34.0	10 2 Q9W910	Q9W910 gb virus c/
25	15	31.9	9 2 Q9W9F4	Q9W9F4 mus musculu
26	15	31.9	9 2 Q8QRR5	Q8QRR5 transmissib
27	15	31.9	9 2 Q8QRR6	Q8QRR6 transmissib
28	15	31.9	10 1 FARP_LOCW1	FARP_LOCW1 locusta mig
29	15	31.9	10 1 LCMS_LEUWA	LCMS_LEUWA leucophaea
30	15	31.9	10 1 NEWS_SARBU	NEWS_SARBU sarcoptaea
31	15	31.9	10 1 XYNB_DICB4	XYNB_DICB4 dictyoglom

32	15	31.9	10 2 Q7S225	Q7S225 neurospora
33	15	31.9	10 2 Q6H101	Q6H101 canis famli
34	15	31.9	10 2 P74843	P74843 streptomyc
35	15	31.9	10 2 P82588	P82588 streptococ
36	15	31.9	10 2 Q9PXC3	Q9PXC3 hepatitis b
37	14	29.8	7 2 Q9YQ10	Q9YQ10 transmissib
38	14	29.8	8 2 Q7Z6G0	Q7Z6G0 homo sapien
39	14	29.8	8 2 Q56246	Q56246 bacillus ps
40	14	29.8	9 1 PPH1_LYCES	PPH1_LYCES
41	14	29.8	9 2 Q7R8X5	Q7R8X5 plasmodium
42	14	29.8	9 2 Q8W5S8	Q8W5S8 anthracidari
43	14	29.8	9 2 P83539	P83539 lactobacilli
44	14	29.8	9 2 Q617X6	Q617X6 macaca neme
45	14	29.8	9 2 Q61754	Q61754 cynops pyrr

## ALIGNMENTS

RESULT 1  
ID Q8X4E5 PRELIMINARY; PRT; 10 AA.  
AC Q8X4E5;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE 25903 protein.  
GN OrderedLocustNames=25903;  
OS Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=83334;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;  
RX MEDLINE=21074935; PubMed=11206551;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D., Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A., Foster G., Hackett J., Klink S., Boutin A., Shao Y., Miller L., Rao Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C., Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";  
RL Nature 409:529-533(2001).  
DR EMBL: AEO05662; AAG59489.1; -;  
DR PIR: E86128; B66128.  
KW Complete proteome.  
SQ SEQUENCE 10 AA; 1161 MW; 98AC9BD2D6D7205A CRC64;  
Query Match 48.9%; Score 23; DB 2; Length 10;  
Best local Similarity 57.1%; Pred. No. 1e+03;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 4 MDLLHV 10  
DB 1 MDLLQV 7  
RESULT 2  
ID Q37854 PRELIMINARY; PRT; 8 AA.  
AC Q37854;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Coliphage genome of unknown function, 5 end. (Fragment).  
OS Bacteriophage K17.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Leviviridae;  
OC Levivirus.  
OX NCBI\_TaxID=12026;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=73224987; PubMed=4352721;

RA Remsing U.F.E.;  
 RT "A sequence of seventy-three nucleotides from the coliphage R17 genome."  
 RL Biochem. J. 131:593-604(1973).  
 DR EMBL: M24820; AAA27255.1; -.  
 FT NON TER 8  
 SQ SEQUENCE 8 AA; 969 MW; ECB45412C1E72726 CRC64;

Query Match 42.6%; Score 20; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LHHV 10  
 DB 2 LHHV 5

## RESULT 3

OS1594 PRELIMINARY; PRT; 8 AA.  
 AC OS1594;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE COB protein (Fragment).  
 OS Plasmid ColV2-K94.  
 OG Plasmid ColV2-K94.  
 OC Plasmids.  
 OX NCBI\_TaxID=2458;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=6623772; PubMed=2423502;  
 RA Weber P.C. Palchaudhuri S.;  
 RT "Incompatibility repressor in a RepA-like replicon of the IncFI plasmid ColV2-K94."  
 RL J. Bacteriol. 166:1106-1121(1986).  
 DR EMBL: M13472; AAA23194.1; -.  
 KW Plasmid.  
 FT NON TER 1  
 SQ SEQUENCE 8 AA; 998 MW; 0F37205AA73416D7 CRC64;

Query Match 38.3%; Score 18; DB 2; Length 8;  
 Best Local Similarity 57.1%; Pred. No. 1.8e+06;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 LAMDL 8  
 DB 1 LORIDL 7

## RESULT 4

RE42\_LITRU STANDARD; PRT; 9 AA.  
 ID RE42\_LITRU  
 AC P82075; P82093;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Rubellidin 4.2/4.3.  
 OS Litoria rubella (Desert tree frog).  
 OC Eukaryota; Chordata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Hylidae;  
 OC Pelodyadinae; Litoria.  
 OX NCBI\_TaxID=104895;  
 RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RC Tissue-Skin secretion;  
 RA Steinboer S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,  
 RA Tyler M.J., Wallace J.C.;  
 RT "The structure of new peptides from the Australian red tree frog 'Litoria rubella'. The skin peptide profile as a probe for the study of evolutionary trends of amphibians."  
 RL Aust. J. Chem. 49:955-963(1996).  
 RN [2]

RP SEQUENCE.  
 RC Tissue-Skin secretion;  
 RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;  
 RT "Peptides from the skin glands of the Australian tree frog Litoria rubella. Comparison with the skin peptides from Litoria rubella."  
 RL Aust. J. Chem. 52:639-645(1999).  
 FT Auct. J. Chem. 52:639-645(1999).  
 SQ SEQUENCE 9 AA; 984 MW; 2C2D7205AA72728 CRC64;

Query Match 38.3%; Score 18; DB 1; Length 9;  
 Best Local Similarity 60.0%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 MLDL 8  
 DB 3 LIDL 7

## RESULT 5

Q7RD8 PRELIMINARY; PRT; 9 AA.  
 AC Q7RD8;  
 DT 01-MAR-2004 (TREMBlrel. 26, Created)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Hypothetical protein.  
 GN Name=PT0516;  
 OS Plasmodium yoelii yoelii.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=73239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=17XNL;  
 RX PubMed=1236865;  
 RA Carlton J.M., Anguoli S.V., Suh B.B., Koc T.W., Pettea M.,  
 RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koc H.L.,  
 RA Peterson J.D., Pop M., Kosack D.S., Skurray M.F., Bidwell S.L.,  
 RA Shalom S.J., van Aken S.E., Riedmiller S.B., Feldblyum T.V.,  
 RA Cho J.K., Quackenbush J., Sedegah M., Shoai A., Cummings L.M.,  
 RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,  
 RA Cunningham D.A., Preiser P.R., Bergman L.W., Valdivia A.B.,  
 RA van Lin H., Jense C.J., Waters A.P., Smith H.O., White O.R.,  
 RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,  
 RA Carucci D.J.;  
 RT "Genome sequence and comparative analysis of the model rodent malaria parasite Plasmodium yoelii yoelii."  
 RL Nature 419:512-519(2002).  
 CC -! CAUTION: The sequence shown here is derived from an EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is preliminary data.  
 DR EMBL: AAB01001760; EAA17550.1; -.  
 KW Hypothetical protein  
 SQ SEQUENCE 9 AA; 1116 MW; 500A3B1337272447 CRC64;

Query Match 38.3%; Score 18; DB 2; Length 9;  
 Best Local Similarity 80.0%; Pred. No. 1.8e+06;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 MLDL 8  
 DB 1 MLDL 5

## RESULT 6

Q7RD8 PRELIMINARY; PRT; 9 AA.  
 AC Q7RD8;  
 DT 01-MAR-2004 (TREMBlrel. 26, Created)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Hypothetical protein.  
 GN Name=PT0516;  
 OS Plasmodium yoelii yoelii.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=73239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=17XNL;  
 RX PubMed=1236865;  
 RA Carlton J.M., Anguoli S.V., Suh B.B., Koc T.W., Pettea M.,  
 RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koc H.L.,  
 RA Peterson J.D., Pop M., Kosack D.S., Skurray M.F., Bidwell S.L.,  
 RA Shalom S.J., van Aken S.E., Riedmiller S.B., Feldblyum T.V.,  
 RA Cho J.K., Quackenbush J., Sedegah M., Shoai A., Cummings L.M.,  
 RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,  
 RA Cunningham D.A., Preiser P.R., Bergman L.W., Valdivia A.B.,  
 RA van Lin H., Jense C.J., Waters A.P., Smith H.O., White O.R.,  
 RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,  
 RA Carucci D.J.;  
 RT "Genome sequence and comparative analysis of the model rodent malaria parasite Plasmodium yoelii yoelii."  
 RL Nature 419:512-519(2002).  
 CC -! CAUTION: The sequence shown here is derived from an EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is preliminary data.  
 DR EMBL: AAB01001760; EAA17550.1; -.  
 KW Hypothetical protein  
 SQ SEQUENCE 9 AA; 1116 MW; 500A3B1337272447 CRC64;

Q66205  
ID 066205 PRELIMINARY; PRT; 7 AA.  
AC 066205;  
DT 01-NOV-1996 (TREMblrel. 01, Created)  
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE Membrane protein (1 is 3rd base in codon) (Fragment).  
OC Transmissible gastroenteritis virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
OC Coronaviridae; Coronavirus; Group 1 species.  
OC NCBI\_TaxID=11149;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ES772/70;  
RA MEDLINE=68216185; PubMed=2835592;  
RA Britton P., Carmenes R.S., Page K.W., Garves D.J., Parra F.;  
RT "Sequence of the nucleoprotein gene from a virulent British field  
RT isolate of transmissible gastroenteritis virus and its expression in  
RT Saccharomyces cerevisiae."  
RL Mol. Microbiol. 2:89-99(1988).  
DR EMBL: Y00542; CAA68606.1; -.  
FT NON TER 1  
SQ SEQUENCE 7 AA; 869 MW; 72C691E727233B70 CRC64;

Query Match 36.2%; Score 17; DB 2; Length 7;  
Best Local Similarity 75.0%; Pred. No. 1.8e+06;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 7 LHM 10  
DB 3 LHM 6

RESULT 7  
Q9TT78 PRELIMINARY; PRT; 8 AA.  
ID 09TT78  
AC 09TT78;  
DT 01-MAY-2000 (TREMblrel. 13, Created)  
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE Thymidylate synthase (Fragment).  
OC Carls Lemiliatis (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedidae; Canidae; Canis.  
OC NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21015404; PubMed=11130975;  
RA Brouillette J.A., Andrew J.R., Venta P.J.;  
RL Mamm. Genome 11:1079-1086(2000).  
DR EMBL: AF202073; AAF20918.1; -.  
FT NON TER 1  
SQ SEQUENCE 8 AA; 899 MW; 6731A1E059CA867 CRC64;

Query Match 34.0%; Score 16; DB 2; Length 8;  
Best Local Similarity 50.0%; Pred. No. 1.8e+06;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 6 DLH 9  
DB 3 DLH 6

RESULT 8  
O70SM2 PRELIMINARY; PRT; 9 AA.  
ID 070SM2  
AC 070SM2;  
DT 05-JUL-2004 (TREMblrel. 27, Created)  
DT 05-JUL-2004 (TREMblrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMblrel. 27, Last annotation update)

DE Hypermethylated in cancer 1 (Fragment).  
GN Name=HIC1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OC NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RA Pinte S., Guerau-del C., Deltour S., Godwin A.R., Leprince D.;  
RT "Identification of a second G-C-rich promoter conserved in the human,  
RT murine and rat tumor suppressor genes HIC1."  
RL Oncogene 23:4023-4031(2004).  
DR EMBL: AF550616; CAD79467.1; -.  
FT NON TER 9  
SQ SEQUENCE 9 AA; 964 MW; 5B5E6DB1681AA7 CRC64;

Query Match 34.0%; Score 16; DB 2; Length 9;  
Best Local Similarity 60.0%; Pred. No. 1.8e+06;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 MDL 8  
DB 1 MDL 5

RESULT 9  
Q8GIZ6 PRELIMINARY; PRT; 9 AA.  
ID 08GIZ6  
AC 08GIZ6;  
DT 01-MAR-2003 (TREMblrel. 23, Created)  
DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)  
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12)  
DE (Fragment).  
GN Name=gap;  
OC Lactobacillus delbrueckii (subsp. lactis).  
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;  
OC Lactobacillus.  
OC NCBI\_TaxID=29397;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCC8;  
RA Bourbiquet A.A., Mollet B.;  
RT "Purification and characterization of the 3-phosphoglycerate kinase  
RT from the thermophile Lactobacillus delbrueckii subsp. lactis."  
RL Int. Dairy J. 12:723-728(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCC8;  
RA Bourbiquet A.A.;  
RT "Molecular insights into the metabolism and physiology of the lactic  
RT acid bacterium Lactobacillus delbrueckii subsp. lactis."  
RL Thesis (2000), Department of Molecular Microbiology, Biozentrum der  
RL Universitaet Basel, (PhD work conducted at the Nestle Research Center,  
RL Lausanne), Switzerland.  
DR EMBL: AJ515554; CAD56494.1; -.  
DR GO: GO:0004365; F:Glyceraldehyde-3-phosphate dehydrogenase (p. .; IEA.  
DR GO: GO:0016491; F:oxidoreductase activity; IEA.  
FT NON TER 1  
SQ SEQUENCE 9 AA; 1071 MW; 94ABADD9C1E72721 CRC64;

Query Match 34.0%; Score 16; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LHM 9  
DB 3 LHM 5

RESULT 10

OSQZ47  
ID 09QZ47 PRELIMINARY; PRT; 9 AA.  
AC 09QZ47.  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
DE TARC protein (Fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Gorski K., Huang X., Tseng S.-Y., Rattis F., Pardoll D., Tsuchiya H.;  
RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF192527; AF04844.1; -  
FT NON TER 9  
SQ SEQUENCE 9 AA; 1104 MW; 621E272686D735A4 CRC64;

Query Match 34.0%; Score 16; DB 2; Length 9;  
Best Local Similarity 80.0%; Pred. No. 1.8e+06;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SLAML 5  
DB 3 SLOWL 7

## RESULT 11

OSUTD7 PRELIMINARY; PRT; 9 AA.  
AC 08UTD7.  
DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
DE Vpu protein.  
GN Name=vpu;  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=21988475; Pubmed=11991972;  
RA Novitsky V., Smith U.R., Gilbert P., McLane M.F., Chigwedere P.,  
RA Williamson C., Ndung'u T., Klein I., Chang S.-Y., Peter T., Thior I.,  
RA Foley B.T., Gaoletxe S., Rybak N., Gaseitlswe S., Vanberg F.,  
RA Marink R., Lee T.-H., Essex M.;  
RT "Human immunodeficiency virus type 1 subtype C molecular phylogeny:  
RT consensus sequence for an AIDS vaccine design";  
RL J. Virol. 76:5435-5451(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Novitsky V.A., McLane M.F., Chigwedere P., Ndung'u T., Klein I.,  
RA Chang S.-Y., Peter T., Thior I., Rybak N., Gaseitlswe S., Vanberg F.,  
RA Marink R., Lee T.-H., Essex M.;  
RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF443091; AL34712.1; -  
SQ SEQUENCE 9 AA; 1102 MW; 188BD40B17272440 CRC64;

Query Match 34.0%; Score 16; DB 2; Length 9;  
Best Local Similarity 60.0%; Pred. No. 1.8e+06;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 MLDL 8  
DB 1 MINL 5

## RESULT 12

CAD79467 PRELIMINARY; PRT; 9 AA.  
ID CAD79467.  
AC CAD79467.  
DT 01-JUN-2004 (TREMBlrel. 27, Created)

DT 01-JUN-2004 (TREMBlrel. 27, Last sequence update)  
DT 01-JUN-2004 (TREMBlrel. 27, Last annotation update)  
DE Hypermethylated in cancer 1 (Fragment).  
GN HIC1.  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Tissue=Mammary gland;  
RX PubMed=15007385;  
RA Pinte S., Guerardel C., Delcours S., Godwin A.K., Lepoint D.;  
RT "Identification of a second G-C-rich promoter conserved in the human,  
RT murine and rat tumor suppressor genes HIC1."  
RL Oncogene 23:4023-4031(2004).  
DR EMBL; AJ550616; CAD79467.1; -  
FT NON TER 9  
SQ SEQUENCE 9 AA; 964 MW; 5B5E6DD81681AA7 CRC64;

Query Match 34.0%; Score 16; DB 2; Length 9;  
Best Local Similarity 60.0%; Pred. No. 1.8e+06;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 MLDL 8  
DB 1 MLDL 5

## RESULT 13

FARP MANSE STANDARD; PRT; 10 AA.  
AC P18523.  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE FMRamide-like neuropeptide.  
DS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).  
OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
CC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Sphingioidea;  
OC Sphingidae; Sphinginae; Manduca.  
OX NCBI\_TaxID=7130;  
RN [1]  
RP SEQUENCE.  
RA MEDLINE=91045350; Pubmed=2235694;  
RA Kingan T.G., Teplov D.B., Phillips J.M., Riehm J.P., Rao K.R.,  
RA Hildebrand J.G., Homberg U., Kammer A.E., Jardine I., Griffin P.R.,  
RA Hunt D.F.;  
RT "A new peptide in the FMRamide family isolated from the CNS of the  
RT hawkmoth, Manduca sexta."  
RL Peptides 11:849-856(1990).  
RN [2]  
RP SEQUENCE.  
RA "FUNCTION: Increases the force of neurally evoked contractions in  
RA the major power-producing flight muscles, the dorsal longitudinal  
RA muscles and so is likely to play a role in sustaining or promoting  
RA flight behavior patterns.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- SIMILARITY: Belongs to the FARP (FMRamide related peptide)  
CC family.  
DR PIR; A43977; A43977.  
KM Amidation: Direct protein sequencing; Neuropeptide;  
KM Pyrrolidone carboxylic acid.  
FT MOD\_RES 10 Phenylalanine amide.  
FT MOD\_RES 10 Pyrrolidone carboxylic acid.  
SQ SEQUENCE 10 AA; 1247 MW; D3C4523D5B1F2D2 CRC64;

Query Match 34.0%; Score 16; DB 1; Length 10;  
Best Local Similarity 50.0%; Pred. No. 2.8e+04;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 6 DLH 9  
DB 2 DVH 5

## RESULT 14

Q9TS43 PRELIMINARY; PRT; 10 AA.  
 AC Q9TS43;  
 DT 01-MAY-2000 (TRENBLREL. 13, Created)  
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)  
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)  
 DE OESTRADIOL-RECEPTOR-PI peptide (fragment).  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9129128; PubMed=2064608;  
 RA Thole H.H., Jungblut P.W., Jakob F.  
 RT "The proton-driven dissociation of oestradiol-receptor dimers as a  
 RT preparative tool. Isolation of a 32 kDa fragment from porcine uteri  
 RT and assignment of C-terminal origin by partial sequencing.";  
 RL Biochem. J. 276:709-714(1991).  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 10 AA; 1240 MW; D31AD274405691F2 CRC64;

## Query Match

Best Local Similarity 34.0%; Score 16; DB 2; Length 10;  
 Best Local Similarity 40.0%; Pred. No. 2.8e+04;  
 Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 6 DLIHIV 10  
 :||:  
 Db 1 ELYTHM 5

## RESULT 15

039949 PRELIMINARY; PRT; 10 AA.  
 ID 039949  
 AC 039949; 039950; 039953; 039954; 039955;  
 DT 01-JAN-1998 (TRENBLREL. 05, Created)  
 DT 01-JAN-1998 (TRENBLREL. 05, Last sequence update)  
 DT 01-NOV-1999 (TRENBLREL. 12, Last annotation update)  
 DE El protein (fragment).  
 OS GB virus C/Hepatitis G virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC GBV-C/HGV group.  
 NCBI\_TaxID=54290;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97368412; PubMed=9225026;  
 RA Smith D.B., Cuccannu N., Davidson P., Jarvis L.M., Mokili J.L.,  
 RA Hamid S., Ludlam C.A., Simmonds P.,  
 RT "Discrimination of hepatitis G virus/GBV-C geographical variants by  
 RT analysis of the 5' non-coding region."  
 RL J. Gen. Virol. 78:1533-1542(1997).  
 DR EMBL; AF003167; AAC57978.1; -  
 FT NON\_TER 10  
 FT NON\_TER 10  
 SQ SEQUENCE 10 AA; 1083 MW; CC88FE27272732 CRC64;

## Query Match

Best Local Similarity 34.0%; Score 16; DB 2; Length 10;  
 Best Local Similarity 55.6%; Pred. No. 2.8e+04;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LAMDLIHY 10  
 :||:  
 Db 1 MAVLLILV 9

Search completed: November 30, 2004, 08:13:39  
 Job time : 80.3333 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:34:52; Search time 55.333 Seconds  
(without alignments)  
64.831 Million cell updates/sec

Title: US-10-008-377A-1  
Perfect score: 47  
Sequence: 1 SLAMDLHV 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 44336

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: A\_Genesep23Sep04:\*

- 1: genesep1980s:\*
- 2: genesep1990s:\*
- 3: genesep2000s:\*
- 4: genesep2001s:\*
- 5: genesep2002s:\*
- 6: genesep2003as:\*
- 7: genesep2003bs:\*
- 8: genesep2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	47	100.0	10 6 AAO22983	AAO22983 Human p28
2	43	91.5	9 6 AAO22985	AAO22985 Human p28
3	39	83.0	9 6 AAO22986	AAO22986 Human p28
4	26	55.3	9 6 ABOU3234	ABOU3234 Human exp
5	26	55.3	10 7 AAW45584	AAW45584 Peptide f
6	26	55.3	10 7 ADR24124	ADR24124 Human 98P
7	26	55.3	10 7 ADR24937	ADR24937 Human 98P
8	25	53.2	9 7 ADR65168	ADR65168 Human VEG
9	25	53.2	10 2 AAR65116	AAR65116 PSA immun
10	25	53.2	10 4 AAG88173	AAG88173 Saccharom
11	25	53.2	10 4 AAG88172	AAG88172 Saccharom
12	24	51.1	9 6 ABR22055	ABR22055 Human can
13	24	51.1	9 6 ABOU3262	ABOU3262 Human exp
14	24	51.1	10 5 AAU94381	AAU94381 Human nov
15	24	51.1	10 7 ADR24903	ADR24903 Human 98P
16	24	51.1	10 8 ADR38377	ADR38377 Human asp
17	23	48.9	9 5 AAE31277	AAE31277 Human mag
18	23	48.9	9 5 AAE31276	AAE31276 Human mag
19	23	48.9	10 2 AAY40366	AAY40366 Amino aci
20	23	48.9	10 2 AAY5317	AAY5317 HIV-1 ne
21	23	48.9	10 2 AAY28658	AAY28658 HIV-derv
22	23	48.9	10 2 ABE97928	ABE97928 CDB+ T ce
23	22.5	47.9	10 5 ABO17292	ABO17292 Zinc tran
24	22.5	47.9	10 5 ABO17024	ABO17024 Zinc tran
25	22	46.8	7 2 AAW46002	AAW46002 Peptide #

26	22	46.8	7 4 AAW98287	AAW98287 Human pep
27	22	46.8	7 8 ADH68228	ADH68228 Human G-p
28	22	46.8	9 2 AAR73813	AAR73813 Antigen f
29	22	46.8	9 2 AAY47722	AAY47722 Immunogen
30	22	46.8	9 2 AAY55465	AAY55465 HLA bindi
31	22	46.8	9 4 AAB75668	AAB75668 HLA class
32	22	46.8	9 5 AAB96083	AAB96083 Human glu
33	22	46.8	9 5 AAE28729	AAE28729 Human GAS
34	22	46.8	9 5 AAB94395	AAB94395 Human nov
35	22	46.8	9 5 AAB94246	AAB94246 Human nov
36	22	46.8	9 5 AAB94858	AAB94858 Human nov
37	22	46.8	9 5 AAB94458	AAB94458 Human nov
38	22	46.8	9 5 AAB15493	AB15493 Zinc tran
39	22	46.8	9 5 AAB16364	AB16364 Zinc tran
40	22	46.8	9 5 AAB16851	AB16851 Zinc tran
41	22	46.8	9 5 AAB15451	AB15451 Zinc tran
42	22	46.8	9 5 AAB15349	AB15349 Zinc tran
43	22	46.8	9 5 AAB15883	AB15883 Zinc tran
44	22	46.8	9 5 AAB15973	AB15973 Zinc tran
45	22	46.8	9 5 AAB16222	AB16222 Zinc tran

## ALIGNMENTS

## RESULT 1

AAO22983 standard; peptide; 10 AA.

XX	AAO22983;	
XX	17-SEP-2003 (first entry)	
DE	Human p280-89 ALK-derived HLA-A*0201 restricted CTL epitope peptide.	
XX	HLA-A*0201 binding; anaplastic lymphoma kinase; ALK; CTL epitope; ALCL;	
XX	cytotoxic T-lymphocyte; receptor tyrosine; NPM/ALK fusion protein; human;	
KM	oncogenic; anaplastic large cell lymphoma; neuroblastoma; gene therapy;	
KM	t(2;5) translocation; neoplasia; cytostatic; immunotherapy; alkoma;	
KM	p280-89.	
OS	Homo sapiens.	
XX	MO2003042243-A2.	
XX	22-MAY-2003.	
XX	14-NOV-2002; 2002WC-EP012764.	
XX	15-NOV-2001; 2001US-00008377.	
PR	(MAST-) IST NAZ STUDIO E CURA DEI TUMORI.	
PA	(SUD-) ST JUDE CHILDREN'S RES HOSPITAL.	
XX	Gambacorti-Passerini C, Passoni L;	
XX	WPI; 2003-441791/41.	
PT	New HLA-A*0201-binding Anaplastic Lymphoma Kinase (ALK) peptide.	
PT	useful for preparing a composition for treating ALK-positive lymphoma;	
PT	neuroblastoma or ALK-expressing neoplasia.	
PS	Claim 1; Page 3; 33p; English.	
CC	The invention relates to a novel HLA-A*0201-binding anaplastic lymphoma	
CC	kinase (ALK)-derived peptide, of 9-10 amino acids, which may act as a	
CC	cytotoxic T-lymphocyte (CTL)-epitope. ALK is a receptor tyrosine kinase	
CC	which forms part of the NPM/ALK fusion protein (oncogenic lymphoma kinase	
CC	fusion protein) in ALCL (anaplastic large cell lymphoma)-derived cell	
CC	lines. More than 50% of ALCL cases possess a t(2;5) chromosomal	
CC	translocation that leads to the expression of the NPM/ALK fusion protein	
CC	which forms a potent oncogene when constitutively activated. Translocated	
CC	ALK is a widely expressed tumour-associated antigen characteristic of ALK	

CC -positive lymphomas, neuroblastomas and ALK-expressing neoplasias. The  
 CC peptides of the invention demonstrate cytostatic activity and induce an  
 CC MHC (major histocompatibility complex) class I restricted cytotoxic  
 CC lymphocyte response against tumour cells expressing the NPM/ALK fusion  
 CC protein. Hence, the peptides may be utilised during the treatment, via  
 CC immunotherapy, of ALK-positive lymphomas such as alkomas, neuroblastomas  
 CC and ALK-expressing neoplasias. In addition, the peptides may be used  
 CC during gene therapy. The current sequence is that of the human p282-89  
 CC ALK-derived HLA-A\*0201 restricted CTL epitope peptide of the invention  
 CC  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 47; DB 6; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMF1LHV 10  
 |||||  
 DB 1 SLAMF1LHV 10

# RESULT 2

AAO22985 AAO22985 standard; peptide; 9 AA.

AC AAO22985;

DT 17-SEP-2003 (first entry)

DE Human p281-89 ALK-derived HLA-A\*0201 restricted CTL epitope peptide.

XX HLA-A\*0201 binding; anaplastic lymphoma kinase; ALK; CTL epitope; ALCL;  
 XX cytotoxic T-lymphocyte; receptor tyrosine; NPM/ALK fusion protein; human;  
 XX oncogenic; anaplastic large cell lymphoma; neuroblastoma; gene therapy;  
 XX t(2;5) translocation; neoplasia; cytostatic; immunotherapy; alkoma;  
 XX p281-89.

KW Homo sapiens.

OS WO2003042243-A2.

PN 22-MAY-2003.

PR 14-NOV-2002; 2002MO-EP012764.

PF 15-NOV-2001; 2001US-00008377.

PI (NAST-) 1ST NAZ STUDIO E CURA DEI TUMORI.

PA (SUD-) ST JUDE CHILDREN'S RES HOSPITAL.

XX Gambacorti-Passerini C, Passoni L;

DR WPI; 2003-441791/41.

PT New HLA-A\*0201-binding Anaplastic Lymphoma Kinase (ALK) peptide,  
 PT useful for preparing a composition for treating ALK-positive lymphoma,  
 PT neuroblastoma or ALK-expressing neoplasia.

PS Claim 1; Page 7; 33pp; English.

XX The invention relates to a novel HLA-A\*0201-binding anaplastic lymphoma  
 CC kinase (ALK)-derived peptide, of 9-10 amino acids, which may act as a  
 CC cytotoxic T-lymphocyte (CTL)-epitope. ALK is a receptor tyrosine kinase  
 CC which forms part of the NPM/ALK fusion protein (oncogenic lymphoma kinase  
 CC fusion protein) in ALCL (anaplastic large cell lymphoma)-derived cell  
 CC lines. More than 50% of ALCL cases possess a t(2;5) chromosomal  
 CC translocation that leads to the expression of the NPM/ALK fusion protein  
 CC which forms a potent oncogene when constitutively activated. Translocated  
 CC ALK is a widely expressed tumour-associated antigen characteristic of ALK  
 CC -positive lymphomas, neuroblastomas and ALK-expressing neoplasias. The  
 CC peptides of the invention demonstrate cytostatic activity and induce an  
 CC MHC (major histocompatibility complex) class I restricted cytotoxic  
 CC lymphocyte response against tumour cells expressing the NPM/ALK fusion

CC protein. Hence, the peptides may be utilised during the treatment, via  
 CC immunotherapy, of ALK-positive lymphomas such as alkomas, neuroblastomas  
 CC and ALK-expressing neoplasias. In addition, the peptides may be used  
 CC during gene therapy. The current sequence is that of the human p281-89  
 CC ALK-derived HLA-A\*0201 restricted CTL epitope peptide of the invention  
 CC  
 SQ Sequence 9 AA;

Query Match 91.5%; Score 43; DB 6; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SLAMF1LHV 10  
 |||||  
 DB 1 SLAMF1LHV 9

# RESULT 3

AAO22986 AAO22986 standard; peptide; 9 AA.

AC AAO22986;

DT 17-SEP-2003 (first entry)

DE Human p282-90 ALK-derived HLA-A\*0201 restricted CTL epitope peptide.

XX HLA-A\*0201 binding; anaplastic lymphoma kinase; ALK; CTL epitope; ALCL;  
 XX cytotoxic T-lymphocyte; receptor tyrosine; NPM/ALK fusion protein; human;  
 XX oncogenic; anaplastic large cell lymphoma; neuroblastoma; gene therapy;  
 XX t(2;5) translocation; neoplasia; cytostatic; immunotherapy; alkoma;  
 XX p282-90.

KW Homo sapiens.

OS WO2003042243-A2.

PN 22-MAY-2003.

PR 14-NOV-2002; 2002MO-EP012764.

PF 15-NOV-2001; 2001US-00008377.

PI (NAST-) 1ST NAZ STUDIO E CURA DEI TUMORI.

PA (SUD-) ST JUDE CHILDREN'S RES HOSPITAL.

XX Gambacorti-Passerini C, Passoni L;

DR WPI; 2003-441791/41.

PT New HLA-A\*0201-binding Anaplastic Lymphoma Kinase (ALK) peptide,  
 PT useful for preparing a composition for treating ALK-positive lymphoma,  
 PT neuroblastoma or ALK-expressing neoplasia.

PS Claim 1; Page 7; 33pp; English.

XX The invention relates to a novel HLA-A\*0201-binding anaplastic lymphoma  
 CC kinase (ALK)-derived peptide, of 9-10 amino acids, which may act as a  
 CC cytotoxic T-lymphocyte (CTL)-epitope. ALK is a receptor tyrosine kinase  
 CC which forms part of the NPM/ALK fusion protein (oncogenic lymphoma kinase  
 CC fusion protein) in ALCL (anaplastic large cell lymphoma)-derived cell  
 CC lines. More than 50% of ALCL cases possess a t(2;5) chromosomal  
 CC translocation that leads to the expression of the NPM/ALK fusion protein  
 CC which forms a potent oncogene when constitutively activated. Translocated  
 CC ALK is a widely expressed tumour-associated antigen characteristic of ALK  
 CC -positive lymphomas, neuroblastomas and ALK-expressing neoplasias. The  
 CC peptides of the invention demonstrate cytostatic activity and induce an  
 CC MHC (major histocompatibility complex) class I restricted cytotoxic  
 CC lymphocyte response against tumour cells expressing the NPM/ALK fusion  
 CC protein. Hence, the peptides may be utilised during the treatment, via  
 CC immunotherapy, of ALK-positive lymphomas such as alkomas, neuroblastomas  
 CC and ALK-expressing neoplasias. In addition, the peptides may be used  
 CC during gene therapy. The current sequence is that of the human p282-90

CC ALK-derived HLA-A\*0201 restricted CTL epitope peptide of the invention  
XX  
SQ Sequence 9 AA;  
Query Match 83.0%; Score 39; DB 6; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 3 AMLDHLHV 10  
DB 1 AMLDHLHV 8  
RESULT 4  
ID ABO03234 standard; protein; 9 AA.  
XX  
AC ABO03234;  
XX  
DT 29-JAN-2003 (first entry)  
XX  
DE Human expressed protein tag (EPT) #14.  
XX  
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO00278524-A2.  
XX  
PD 10-OCT-2002.  
XX  
PF 28-MAR-2002; 2002WO-US009671.  
XX  
PR 28-MAR-2001; 2001US-0279495P.  
PR 21-MAY-2001; 2001US-0292544P.  
PR 08-AUG-2001; 2001US-0310801P.  
PR 01-OCT-2001; 2001US-0326370P.  
PR 04-DEC-2001; 2001US-0336780P.  
PR 20-FEB-2002; 2002US-0358985P.  
XX  
PA (ZYCO-) ZYCO INC.  
XX  
PI Chicx RM, Tomlinson AU, Urban RG;  
XX  
DR WPI; 2003-040607/03.  
XX  
PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
PT cytoskeletal proteins, receptors or transcription factors), useful for  
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
PT leukemia.  
XX  
PS Claim 10; SEQ ID NO 14; 134pp; English.  
XX  
CC The invention describes a purified polypeptide, which comprises a  
CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
CC transporter, cytoskeletal protein, receptor or transcription factor. The  
CC polypeptide is useful as an immunogenic composition for eliciting in a  
CC mammal an immunogenic response directed against any of the purified  
CC polypeptide. The purified polypeptide, or the antibody that binds to this  
CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WPI at

CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 9 AA;  
Query Match 55.3%; Score 26; DB 6; Length 9;  
Best Local Similarity 71.4%; Pred. No. 1.7e+06;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
CY 3 AMLDHLHV 9  
DB 1 ALDHLHV 7  
RESULT 5  
ID AAW45584 standard; peptide; 10 AA.  
XX  
AC AAW45584;  
XX  
DT 03-JUN-1998 (first entry)  
XX  
DE Peptide fragment of leptin (ob 106-115) that modulates body weight.  
XX  
KW Leptin; obesity; body weight; diabetes; energy; metabolic disorder;  
KW ob protein.  
XX  
OS Homo sapiens.  
XX  
PN WO9746585-A2.  
XX  
PD 11-DEC-1997.  
XX  
PF 04-JUN-1997; 97WO-EP02968.  
XX  
PR 06-JUN-1996; 96GB-00011775.  
PR 05-SEP-1996; 96GB-00018540.  
PR 20-FEB-1997; 97GB-00003493.  
XX  
PA (SMIK ) SMITHKLINE BEECHAM PLC.  
XX  
PI Albarazani KA, Arch JR, Camilleri P, Neville WA;  
XX  
DR WPI; 1998-042120/04.  
XX  
PT Peptide fragments of leptin that modulate body weight by regulating  
PT energy utilisation - especially useful for treatment of obesity and  
PT diabetes.  
XX  
PS Claim 4; Page 1; 19pp; English.  
XX  
CC The invention relates to specifically claimed peptides AAW45577-W45586 or  
CC their derivatives, analogues and variants, that modulate, specifically  
CC reduce, body weight, mainly by affecting energy utilisation. Also new  
CC are: (1) nucleic acid that encodes the peptides; (2) vectors containing  
CC the nucleic acid; and (3) host cells transformed with this vector. The  
CC peptides are used to treat nutritional or metabolic disorders,  
CC particularly obesity and diabetes  
XX  
SQ Sequence 10 AA;  
Query Match 55.3%; Score 26; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 6 DLHLHV 10  
DB 1 DLHLHV 5  
RESULT 6  
ID ADK24124 standard; peptide; 10 AA.  
XX

```

AC   ADRK4124.4
AD   06-MAY-2004 (first entry)
DE   Human 98P4B6 derived motif-bearing CTL peptide epitope #431.
XX
XX
XX   human; 98P4B6; STEP-2; cancer; breast; colon; lung; kidney; prostate;
XX   gene therapy; genetic abnormality; transgenic; knockout animal;
XX   cystocytic; epitope.
XX
XX   Homo sapiens.
XX
XX   WC2003087306-AA2.
XX
XX   23-OCT-2003.
XX
XX
XX   04-APR-2003; 2003WO-US010462.
XX
XX   05-APR-2002; 2002US-03703877.
XX   06-JUN-2002; 2002US-00165044.
XX   20-DEC-2002; 2002US-0435480P.
XX
XX   (AGEN-) AGENYSYS INC.
XX
XX   Chailita-Eid PM, Raltano AB, Faris M, Ge W, Jakobovits A;
XX   WPI; 2003-903158/32.
XX
XX
XX   A composition comprising 98P4B6 proteins and nucleic acid molecules for
XX   detecting, preventing, prognosing and/or treating cancers that express
XX   98P4B6 (e.g. pancreatic, prostatic or lung cancer).
XX
XX   Claim 1; Page 146; 616pp; English.
XX
XX
XX   This invention relates to a novel human gene 98P4B6 (also known as STEP-
XX   2), six transmembrane epithelial antigen of prostate-2) located on
XX   chromosome 7q21 and the encoded protein and variants derived thereof.
XX   Specifically, it refers to the expression pattern of this gene in adult
XX   normal tissues and its aberrant over-expression in various cancers
XX   including breast, colon, lung, kidney and prostate. The present invention
XX   describes compositions and methods useful for detecting, preventing,
XX   prognosing and/or treating cancers that express 98P4B6. Furthermore, the
XX   polynucleotide can be used for gene therapy purposes, for monitoring
XX   genetic abnormalities and for generating transgenic or 'knockout'
XX   animals, which can be useful for the development and screening of
XX   therapeutically useful reagents. The encoded proteins may also be used in
XX   generating antibodies and vaccines, as well as in identifying ligands and
XX   cellular constituents that bind to 98P4B6 gene products. Accordingly,
XX   these compositions exhibit cytostatic activities. This peptide sequence
XX   is a motif bearing CTL peptide epitope derived from a human 98P4B6
XX   protein variant of the invention.
XX
XX   Sequence 10 AA;
XX
XX   Query Match          55.3%; Score 26; DB 7; Length 10;
XX   Best Local Similarity 50.0%; Pred. No. 1.8e+02;
XX   Matches      5; Conservative      3; Mismatches      2; Gaps      0;
XX
XX   1 SIAMDLIHY 10
XX   |:::|||||*
XX   1 SIYLDLQL 10
XX
XX
XX   RESULT 7
XX   ADRK2937
XX   ADRK4937 standard; peptide; 10 AA.
XX
XX   ADRK4937;
XX
XX   06-MAY-2004 (first entry)
XX
XX   Human 98P4B6 derived motif-bearing CTL peptide epitope #1047.
XX

```

XX	human_98P486; STRAP-2; cancer; breast; colon; lung; kidney; prostate;
KW	gene therapy; genetic abnormality; transgenic; knockout animal;
XX	cytostatic; epitope.
XX	
OS	Homo sapiens.
XX	
FN	WO2003087306-A2.
XX	
PD	23-OCT-2003.
XX	
PF	04-APR-2003; 2003WO-US010462.
XX	
PR	05-APR-2002; 2002US-0370387P.
XX	
PR	06-JUN-2002; 2002US-0016504P.
XX	
PR	20-DEC-2002; 2002US-0435480P.
XX	
PA	(AGEN-) AGENSYS INC.
XX	
PI	Challita-Eld PM, Raitano AB, Faris M, Ge W, Jakobovits A;
XX	
DR	WPI; 2003-903158/82.
XX	
PT	A composition comprising 98P486 proteins and nucleic acid molecules for
XX	detecting, preventing, prognosis and/or treating cancers that express
PT	98P486 (e.g. pancreatic, prostatic or lung cancer).
XX	
PS	Claim 1, Page 152; 616pp; English.
XX	
CC	This invention relates to a novel human gene 98P486 (also known as STRAP-
XX	2, six transmembrane epithelial antigen of prostate-2) located on
CC	chromosome 7q21 and the encoded protein and variants derived thereof.
XX	Specifically, it refers to the expression pattern of this gene in adult
CC	normal tissues and its aberrant over-expression in various cancers
XX	including breast, colon, lung, kidney and prostate. The present invention
CC	describes compositions and methods useful for detecting, preventing,
XX	prognosis and/or treating cancers that express 98P486. Furthermore, the
CC	polynucleotide can be used for gene therapy purposes, for monitoring
XX	genetic abnormalities and for generating transgenic or 'knockout'
CC	animals, which can be useful for the development and screening of
XX	therapeutically useful reagents. The encoded proteins may also be used in
CC	generating antibodies and vaccines, as well as in identifying ligands and
XX	cellular constituents that bind to 98P486 gene products. Accordingly,
CC	these compositions exhibit cytostatic activities. This peptide sequence
XX	is a motif bearing CTL peptide epitope derived from a human 98P486
CC	protein variant of the invention.
XX	
SO	Sequence 10 AA:
XX	
Query Match	55.3%; Score 26; DB 7; Length 10;
Best Local Similarity	50.0%; Pred. NO. 1.8e+02;
Matches	5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY	1 SLAMDDLHV 10
DB	1 SLVLDLQL 10
XX	
RESULT 8	
ADK65168	ID ADK65168 standard; peptide; 9 AA.
XX	
ADK65168;	
XX	
DT	06-MAY-2004 (first entry)
XX	
DE	Human VEGF-A-derived peptide #14 for anti-angiogenesis treatment.
XX	
XX	antiangiogenic; cytostatic; antiinflammatory; immunosuppressive;
XX	ophthalmological; gynecological; antiarteriosclerotic; vitreous;
KW	heparotrophic; dermatological; anti-HIV; antidiabetic; antiproliferative;
XX	antiinflammatory; antiarthritic; antihypertensive; immunization; angiogenesis;
KW	vascular endothelial growth factor; VEGF; neuropeptide; inflammation;
XX	placental growth factor; tumor; neoplasias; metastases;

KW autoimmunity; eye disease; arthritis; endometriosis; arteriosclerosis;  
 KW edema; hepatitis; Kaposi sarcoma; diabetes; psoriasis;  
 KW rheumatoid arthritis; thyroiditis; diabetic retinopathy;  
 KW transplant rejection; macular degeneration; neovascular glaucoma;  
 KW hemangioma; angiodioma.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003086450-A1.  
 XX  
 PD 23-OCT-2003.  
 XX  
 PF 11-APR-2003; 2003WO-CU000004.  
 XX  
 PR 15-APR-2002; 2002CU-00000076.  
 XX  
 PA (INGG-) CENT ING GENETICA & BIOTECNOLOGIA.  
 XX  
 PI Bequet Romero M, Acevedo Castro BE, Gavilondo Cowley JV,  
 PI Fernandez Molina LE, Lopez Ochoa O, Silva Rodriguez RDC,  
 PI Musachio Iasa A, Galban Rodriguez E, Vazquez Blomquist DM;  
 XX  
 DR WPI; 2003-833615/77.  
 XX  
 PT Active immunization against angiogenic proteins, useful for treating e.g.  
 PT tumors and inflammation, particularly contains vascular endothelial  
 PT growth factor or its receptor.  
 PS  
 PS Disclosure; Page 18; 53pp; Spanish.  
 XX  
 CC The invention relates to an active immunization against angiogenic  
 CC proteins comprising administration of a vaccination composition (A),  
 CC optionally containing an adjuvant, that comprises polypeptides (I),  
 CC directly associated with an increase in angiogenesis, their variants, or  
 CC their encoding polynucleotides (II). Angiogenesis-associated polypeptides  
 CC are: members of the vascular endothelial growth factor (VEGF) family,  
 CC especially the 121, 165 or 189 isoforms of VEGF-A, the 167 isoform of  
 CC VEGF-B, or VEGF-C or -D; a (co-)receptor of VEGF, particularly VEGFR-1,  
 CC 2 or -3; NRP-1 or -2 (members of the neuropilin family); or placental  
 CC growth factor. (A) is used for treatment or prevention of tumors in  
 CC mammals, particularly humans but also farm animals and pets, also many  
 CC other conditions associated with excessive angiogenesis, specifically  
 CC malignant or benign neoplasias (and their metastases), acute or chronic  
 CC inflammation, autoimmunity and eye diseases (claimed). Among the diseases  
 CC that may be treated are arthritis, endometriosis, arteriosclerosis,  
 CC edema, infectious diseases (hepatitis and Kaposi sarcoma), diabetes,  
 CC psoriasis, rheumatoid arthritis, thyroiditis, diabetic retinopathy,  
 CC transplant rejection, macular degeneration, neovascular glaucoma,  
 CC hemangioma and angiodioma. The method destroys cells that are the  
 CC source of angiogenic proteins, rather than just neutralizing the activity  
 CC of such proteins (as in passive immunization). This sequence represents  
 CC an immunisation peptide of the invention derived from the VEGF proteins.  
 CC  
 XX  
 SQ Sequence 9 AA;  
 XX  
 Query Match 53.2%; Score 25; DB 7; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1.7e+06;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 OY 1 SLAMDLIH 9  
 DB 1 SLALILYH 9  
 XX  
 RESULT 9  
 AAR65116  
 ID AAR65116 standard; peptide; 10 AA.  
 XX  
 AC AAR65116;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 09-OCT-1995 (first entry)  
 XX

DE PSA immunogenic peptide 166-175.  
 XX  
 XX PSA; immunogenic peptide 166-175; cytotoxic T cells; in vitro activation;  
 KW cancer; AIDS; bacterial infections; malaria; fungal infections;  
 KW tuberculosis; hepatitis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9504817-A1.  
 XX  
 PD 16-FEB-1995.  
 XX  
 PF 01-AUG-1994; 94WO-US008672.  
 XX  
 PR 06-AUG-1993; 93US-00103401.  
 XX  
 PA (CYTE-) CYTEL CORP.  
 XX  
 PI Celis E, Kubo R, Serra H, Tsai V, Wentworth P,  
 PI WPI; 1995-090895/12.  
 XX  
 DR In vitro activation of cytotoxic T cells for selected killing of target  
 XX cells - for treating e.g. cancer, AIDS, hepatitis etc. by incubating them  
 XX with antigen presenting cells loaded with appropriate immunogenic  
 XX peptide.  
 XX  
 PS Example 3; Page 35; 53pp; English.  
 XX  
 CC AAR65109-R65145 are immunogenic peptides, they are used in a new method  
 CC for the in vitro activation of cytotoxic T cells (CTC). This is achieved  
 CC by incubating the CTCs with antigen presenting cells loaded with an  
 CC appropriate immunogenic peptide (e.g. one of the above peptides). By  
 CC selecting the peptides used the following diseases and infections can be  
 CC treated; cancer, AIDS, hepatitis, other viral and bacterial infections,  
 CC malaria and tuberculosis. (Updated on 25-MAR-2003 to correct PN field.)  
 CC  
 XX  
 SQ Sequence 10 AA;  
 XX  
 Query Match 53.2%; Score 25; DB 2; Length 10;  
 Best Local Similarity 44.4%; Pred. No. 2.9e+02;  
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 OY 2 LAMDLIHV 10  
 DB 2 LQCVDLVHI 10  
 XX  
 RESULT 10  
 AAG88173  
 ID AAG88173 standard; peptide; 10 AA.  
 XX  
 AC AAG88173;  
 XX  
 DT 11-SEP-2001 (first entry)  
 DT Saccharomyces cerevisiae peptide; SEQ ID NO: 3122.  
 DE Saccharomyces cerevisiae; complementary peptide; peptide identification;  
 KW Saccharomyces cerevisiae; complementary peptide; peptide identification;  
 KW drug discovery; drug design.  
 XX  
 OS Saccharomyces cerevisiae.  
 XX  
 PN WO200142276-A1.  
 XX  
 PD 14-JUN-2001.  
 XX  
 PF 13-DEC-2000; 2000WO-GB004773.  
 XX  
 PR 13-DEC-1999; 99GB-00028471.  
 XX  
 PA (PROT-) PROTEOM LTD.  
 XX

PI Roberts GW, Heal JR;  
 XX WPI; 2001-367863/38.  
 XX  
 XX Identifying complementary peptides by analysis of protein and nucleotide  
 PT sequence databases, useful in drug design.  
 XX  
 XX Example 5; Page 460; 488pp; English.  
 XX  
 CC The invention relates to the identification of complementary peptides by  
 CC analysis of protein and nucleotide sequence databases from higher  
 CC eukaryotic genomes, excluding human and plants. The specific  
 CC complementary peptides interact with their relevant target proteins  
 CC encoded in the eukaryote genome. The peptides may be used as reagents and  
 CC drugs for drug discovery and as lead ligands for drug design and  
 CC development. The present sequence is a complementary peptide from  
 CC Saccharomyces cerevisiae  
 XX  
 SQ Sequence 10 AA;  
 Query Match 53.2%; Score 25; DB 4; Length 10;  
 Best Local Similarity 62.5%; Pred. No. 2.9e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 SLAMDILH 9  
 Db 1 LGMKELH 8  
 RESULT 11  
 ID AAG8172  
 XX AAG8172 standard; peptide; 10 AA.  
 XX  
 AC AAG8172;  
 XX  
 DT 11-SEP-2001 (first entry)  
 XX  
 DE Saccharomyces cerevisiae peptide, SEQ ID NO: 3121.  
 XX  
 KM Saccharomyces cerevisiae; complementary peptide; peptide identification;  
 KM drug discovery; drug design.  
 XX  
 OS Saccharomyces cerevisiae.  
 XX  
 PN WO200142276-A1.  
 XX  
 PD 14-JUN-2001.  
 XX  
 PF 13-DEC-2000; 2000WO-GB004773.  
 XX  
 PR 13-DEC-1999; 99GB-00029471.  
 XX  
 PA (PROT-) PROTEOM LTD.  
 XX  
 PI Roberts GW, Heal JR;  
 XX  
 DR WPI; 2001-367863/38.  
 XX  
 PT Identifying complementary peptides by analysis of protein and nucleotide  
 PT sequence databases, useful in drug design.  
 XX  
 XX Example 5; Page 460; 488pp; English.  
 XX  
 CC The invention relates to the identification of complementary peptides by  
 CC analysis of protein and nucleotide sequence databases from higher  
 CC eukaryotic genomes, excluding human and plants. The specific  
 CC complementary peptides interact with their relevant target proteins  
 CC encoded in the eukaryote genome. The peptides may be used as reagents and  
 CC drugs for drug discovery and as lead ligands for drug design and  
 CC development. The present sequence is a complementary peptide from  
 CC Saccharomyces cerevisiae  
 XX  
 SQ Sequence 10 AA;

Query Match 53.2%; Score 25; DB 4; Length 10;  
 Best Local Similarity 62.5%; Pred. No. 2.9e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 SLAMDILH 9  
 Db 1 LGMKELH 8  
 RESULT 12  
 ID ABR22055  
 XX ABR22055 standard; peptide; 9 AA.  
 XX  
 AC ABR22055;  
 XX  
 DT 19-MAY-2003 (first entry)  
 XX  
 DE Human cancer-related protein 185P2C9 HLA peptide #1890.  
 XX  
 KM Human, cytostatic; vaccine; cancer; immune response; HLA;  
 KM human leukocyte antigen.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200283921-A2.  
 XX  
 PD 24-OCT-2002.  
 XX  
 PF 10-APR-2002; 2002WO-US011654.  
 XX  
 PR 10-APR-2001; 2001US-0282739P.  
 XX  
 PR 10-APR-2001; 2001US-0283112P.  
 XX  
 PR 25-APR-2001; 2001US-0286630P.  
 XX  
 PA (AGEN-) AGENSYS INC.  
 XX  
 PI Jakobovits A, Chailita-Bid PM, Faris M, Ge W, Hubert RS;  
 PI Morrison K, Morrison RK, Raitano AB;  
 XX  
 DR WPI; 2003-075555/07.  
 XX  
 PT New composition comprising a substance that modulates the structure of  
 PT proteins and polynucleotides, useful for therapeutic, prognostic and  
 PT diagnostic reagents for eliciting cellular or humoral immune response in  
 PT cancer patients.  
 XX  
 PS Claim 13; Page 341; 1021pp; English.  
 XX  
 CC The present invention relates to novel human cancer-related genes and  
 CC proteins (AB278120-AB278168 and ABR01789-ABR01861). The genes and  
 CC proteins are useful for eliciting a humoral or cellular immune response.  
 CC The genes are useful as probes and primers for the amplification and/or  
 CC detection of genes, mRNAs or their fragments, as reagents for the  
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of  
 CC directing the expression of the protein, as tools for modulating or  
 CC inhibiting the expression of genes and/or translation of transcripts, and  
 CC as therapeutic agents. The proteins and peptides are useful as  
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present  
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example  
 CC from the invention  
 XX  
 SQ Sequence 9 AA;  
 Query Match 51.1%; Score 24; DB 6; Length 9;  
 Best Local Similarity 55.6%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 SLAMDILH 9  
 Db 1 SLAMDILH 9

## RESULT 13

ABU03262 standard; protein; 9 AA.

ABU03262;

29-JAN-2003 (first entry)

Human expressed protein tag (EPT) #42.

Translational profiling; expressed protein tag; EPT; kinase; phosphatase; proteinase; protease inhibitor; transporter; cytoskeletal protein; receptor; transcription factor; cancer; MHC; major histocompatibility complex; myeloma; colon cancer; gastric cancer; adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.

Homo sapiens.

WO200278524-A2.

10-OCT-2002.

28-MAR-2002; 2002WO-US009671.

28-MAR-2001; 2001US-0279495P.

21-MAY-2001; 2001US-0292544P.

08-AUG-2001; 2001US-0310801P.

01-OCT-2001; 2001US-0326370P.

04-DEC-2001; 2001US-0336780P.

20-FEB-2002; 2002US-0358985P.

(ZYCO-) ZYCO INC.

Chicz RM, Tomlinson AJ, Urban RG;

WPI; 2003-040607/03.

New polypeptides (e.g. kinases, phosphatases, proteases, transporters, cytoskeletal proteins, receptors or transcription factors), useful for treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or leukemia.

Claim 10; SEQ ID NO 42; 134pp; English.

The invention describes a purified polypeptide, which comprises a fragment of a kinase, phosphatase, protease or transcription factor. The polypeptide is useful as an immunogenic composition for eliciting in a mammal an immunogenic response directed against any of the purified polypeptide. The purified polypeptide, or the antibody that binds to this polypeptide, is useful for treating cancer. The polypeptide is also useful for identifying compounds that binds to a naturally processed class I or class II MHC-binding polypeptide. The polypeptides and polynucleotides are particularly useful for treating or preventing myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma, lymphoma or leukemia. These are also useful for screening agents for treating the above mentioned diseases. This sequence represents an expressed protein tag (EPT) isolated from human tissue for translational profiling. Note: This sequence does not appear in the printed specification but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences)

Sequence 9 AA;

Query Match 51.1%; Score 24; DB 6; Length 9;

Best Local Similarity 50.0%; Pred. No. 1.7e+06; Mismatches 0; Indels 0; Gaps 0;

Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 MDDLMH 9  
|:|:|  
DB 2 MVDIHH 7

## RESULT 14

AAU94381 standard; peptide; 10 AA.

AAU94381;

02-JUL-2002 (first entry)

Human novel protein CatrF2E11 HLA binding peptide #164.

Human; human leukocyte antigen; HLA; immunogen; 83P2H3; CatrF2E11; calcium transport protein; cancer; prostate cancer; cytostatic; chromosome 7q34; chromosome 12q24.1; T cell, B cell.

Homo sapiens.

WO200214361-A2.

21-FEB-2002.

17-AUG-2001; 2001WO-US025782.

17-AUG-2000; 2000US-0226329P.

(AGEN-) AGENSYS INC.

Raitano AB, Chailita-Elid PM, Faris M, Saffran DC, Afar DEH;

Levin E, Hubert RS, Ge W, Jakobovits A;

WPI; 2002-269179/31.

Monitoring 83P2H3 gene products for monitoring the presence of cancer in a subject, comprises determining the status of 83P2H3 gene products in a tissue sample from the subject and comparing it to a normal sample.

Example 11; Page 176; 270pp; English.

The invention relates to monitoring 83P2H3 (a calcium transport protein whose gene is located on chromosome 7q34) gene products in a biological sample from a patient who has or is suspected of having cancer (especially prostate cancer), comprising: (a) determining the status of 83P2H3 gene products expressed by cells in a tissue sample from an individual and (b) comparing the status of 83P2H3 gene products in a normal sample. Also included are modulators of 83P2H3 function or status, generating antibodies/immune response against 83P2H3 (or related protein CatrF2E11 whose gene is located on chromosome 12q24.1) using identified HLA (human leukocyte antigen) binding peptides derived from the protein, delivering a cytotoxic agent to a cell expressing 83P2H3 by conjugating the agent to an anti-83P2H3 antibody, a recombinant protein comprising an antigen-binding region of the antibody, a non-human transgenic animal that produces the recombinant protein, a hybridoma that produces the recombinant protein, a single-chain monoclonal antibody that comprises the variable domains of the heavy and light chains of the anti-83P2H3 antibody, a vector comprising a polynucleotide that encodes the monoclonal antibody and inducing an immune response to a 83P2H3 protein, by providing a 83P2H3-related protein that comprises a T cell or B cell epitope, and contacting the epitope with an immune system T cell or B cell, respectively. The method is useful for monitoring 83P2H3 gene products in a biological sample for monitoring the presence of cancer in an individual. The modulator is useful for inhibiting the growth of cancer cells that express 83P2H3, for treating cancer and the vector is useful for treating a patient with a cancer that expresses 83P2H3. The immunological methods are useful for generating an immune response against 83P2H3, and for detecting the presence of 83P2H3-related protein or polynucleotide in a biological sample from a patient who has or is suspected of having cancer. The antibody is useful in prostate cancer diagnosis, prognosis, imaging methodologies and treatment, to detect and quantify 83P2H3 and mutant 83P2H3-related proteins, for purifying a 83P2H3-related protein, for isolating 83P2H3 homologues/related molecules, and for generating anti-idiotypic antibodies that mimic the 83P2H3 protein. The present sequence is an HLA binding peptide motif from 83P2H3 or its related protein CatrF2E11

XX Sequence 10 AA;

Query Match 51.1%; Score 24; DB 5; Length 10;  
 Best Local Similarity 66.7%; Pred. No. 4.5e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 LAMLDLHV 10  
 |||  
 2 LAALLLVHV 10

# RESULT 15

ID ADK24903 standard; peptide; 10 AA.

AC ADK24903;

DT 06-MAY-2004 (first entry)

DE Human 98P4B6 derived motif-bearing CTL peptide epitope #1013.

KM human; 98P4B6; STEAP-2; cancer; breast; colon; lung; kidney; prostate;  
 KW gene therapy; genetic abnormality; transgenic; knockout animal;  
 CY cytosolic; epitope.

OS Homo sapiens.

FN W02003087306-A2.

PD 23-OCT-2003.

PF 04-APR-2003; 2003NO-US010462.

PR 05-APR-2002; 2002US-0370387P.

PR 06-JUN-2002; 2002US-00165044.

PR 20-DEC-2002; 2002US-0435480P.

PA (AGEN-) AGENSYS INC.

PI Chailita-Eid PM, Raitano AB, Faris M, Ge W, Jakobovits A;

DR WPI; 2003-903158/82.

PT A composition comprising 98P4B6 proteins and nucleic acid molecules for  
 detecting, preventing, prognosing and/or treating cancers that express  
 98P4B6 (e.g. pancreatic, prostatic or lung cancer).  
 PS Claim 1; Page 152; 616p; English.

CC This invention relates to a novel human gene 98P4B6 (also known as STEAP-  
 2, six transmembrane epithelial antigen of prostate-2) located on  
 CC chromosome 7q21 and the encoded protein and variants derived thereof.  
 CC Specifically, it refers to the expression pattern of this gene in adult  
 CC normal tissues and its aberrant over-expression in various cancers  
 CC including breast, colon, lung, kidney and prostate. The present invention  
 CC describes compositions and methods useful for detecting, preventing,  
 CC prognosing and/or treating cancers that express 98P4B6. Furthermore, the  
 CC polynucleotide can be used for gene therapy purposes, for monitoring  
 CC genetic abnormalities and for generating transgenic or 'knockout'  
 CC animals, which can be useful for the development and screening of  
 CC therapeutically useful reagents. The encoded proteins may also be used in  
 CC generating antibodies and vaccines, as well as in identifying ligands and  
 CC cellular constituents that bind to 98P4B6 gene products. Accordingly,  
 CC these compositions exhibit cytostatic activities. This peptide sequence  
 CC is a motif bearing CTL peptide epitope derived from a human 98P4B6  
 CC protein variant of the invention.

XX Sequence 10 AA;

Query Match 51.1%; Score 24; DB 7; Length 10;  
 Best Local Similarity 60.0%; Pred. No. 4.5e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 SLAMDLHV 10  
 |||  
 1 SLGLSLAV 10

Search completed: November 30, 2004, 08:09:32  
 Job time : 57.333 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 08:13:53 ; Search time 44 Seconds

(without alignments)  
80.612 Million cell updates/sec

Title: US-10-008-377a-1

Perfect score: 47

Sequence: 1 SLAMDILHV 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1575965 seqs, 354694765 residues

Total number of hits satisfying chosen parameters: 177356

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10D\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10D\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	100.0	10	US-10-008-377-1	Sequence 1, Appli
2	43	91.5	9	US-10-008-377-3	Sequence 3, Appli
3	39	83.0	9	US-10-008-377-4	Sequence 4, Appli
4	26	55.3	10	US-09-844-774-9	Sequence 9, Appli
5	26	55.3	10	US-10-290-144-9	Sequence 9, Appli
6	24	51.1	8	US-10-064-903-2	Sequence 2, Appli
7	24	51.1	10	US-09-932-165-364	Sequence 364, App
8	22.5	47.9	10	US-10-024-652-1548	Sequence 1548, Ap
9	22.5	47.9	10	US-10-024-652-1816	Sequence 1816, Ap
10	22	46.8	7	US-09-758-128-30	Sequence 30, Appli
11	22	46.8	7	US-09-758-426-30	Sequence 30, Appli
12	22	46.8	7	US-09-758-198-30	Sequence 30, Appli
13	22	46.8	7	US-09-861-661-30	Sequence 30, Appli

14	22	46.8	7	US-10-400-991-27	Sequence 27, Appli
15	22	46.8	8	US-10-367-580-170	Sequence 170, App
16	22	46.8	8	US-10-367-593-170	Sequence 170, App
17	22	46.8	8	US-10-367-594-170	Sequence 170, App
18	22	46.8	8	US-10-367-654-170	Sequence 170, App
19	22	46.8	8	US-10-367-658-170	Sequence 170, App
20	22	46.8	8	US-10-367-658-170	Sequence 170, App
21	22	46.8	8	US-10-367-674-170	Sequence 170, App
22	22	46.8	8	US-09-393-634-88	Sequence 88, Appli
23	22	46.8	9	US-09-753-126-119	Sequence 119, Appli
24	22	46.8	9	US-09-886-896A-83	Sequence 83, Appli
25	22	46.8	9	US-09-932-165-223	Sequence 223, App
26	22	46.8	9	US-09-932-165-841	Sequence 441, App
27	22	46.8	9	US-09-932-165-841	Sequence 841, App
28	22	46.8	9	US-09-932-165-868	Sequence 968, App
29	22	46.8	9	US-10-383-982-88	Sequence 88, Appli
30	22	46.8	9	US-10-334-726-157	Sequence 157, App
31	22	46.8	9	US-10-024-652-3	Sequence 3, Appli
32	22	46.8	9	US-10-024-652-105	Sequence 105, App
33	22	46.8	9	US-10-024-652-147	Sequence 147, App
34	22	46.8	9	US-10-024-652-236	Sequence 236, App
35	22	46.8	9	US-10-024-652-442	Sequence 442, App
36	22	46.8	9	US-10-024-652-510	Sequence 510, App
37	22	46.8	9	US-10-024-652-537	Sequence 537, App
38	22	46.8	9	US-10-024-652-627	Sequence 627, App
39	22	46.8	9	US-10-024-652-642	Sequence 642, App
40	22	46.8	9	US-10-024-652-868	Sequence 868, App
41	22	46.8	9	US-10-024-652-840	Sequence 840, App
42	22	46.8	9	US-10-024-652-1046	Sequence 1046, App
43	22	46.8	9	US-10-024-652-1110	Sequence 1110, App
44	22	46.8	9	US-10-024-652-1111	Sequence 1111, App
45	22	46.8	9	US-10-024-652-1201	Sequence 1201, App

## ALIGNMENTS

RESULT 1  
US-10-008-377-1  
; Sequence 1, Application US/10008377  
; Publication No. US20030157101A1  
; GENERAL INFORMATION:  
; APPLICANT: Gambacorti-Passerini, Carlo  
; TITLE OF INVENTION: Immunogenic ALK Peptides  
; FILE REFERENCE: 045922/241203  
; CURRENT APPLICATION NUMBER: US/10/008,377  
; CURRENT FILING DATE: 2001-11-15  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-008-377-1

Query Match 100.0% Score 47; DB 14; Length 10;  
Best Local Similarity 100.0% Pred. No. 0.066;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMDILHV 10  
DB 1 SLAMDILHV 10

RESULT 2  
US-10-008-377-3  
; Sequence 3, Application US/10008377  
; Publication No. US20030157101A1  
; GENERAL INFORMATION:  
; APPLICANT: Gambacorti-Passerini, Carlo  
; APPLICANT: Passoni, Lorena  
; TITLE OF INVENTION: Immunogenic ALK Peptides

FILE REFERENCE: 045922/241203  
CURRENT APPLICATION NUMBER: US/10/008,377  
CURRENT FILING DATE: 2001-11-15  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-008-377-3

Query Match  
Best Local Similarity 91.5%; Score 43; DB 14; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0;

QY 2 LAMIDLHV 10  
DB 1 LAMIDLHV 9

RESULT 3  
US-10-008-377-4  
Sequence 4, Application US/10008377  
Publication No. US20030157101A1  
GENERAL INFORMATION:  
APPLICANT: Gambacorti-Passerini, Carlo  
APPLICANT: Passoni, Lorena  
TITLE OF INVENTION: Immunogenic ALK Peptides  
FILE REFERENCE: 045922/241203  
CURRENT APPLICATION NUMBER: US/10/008,377  
CURRENT FILING DATE: 2001-11-15  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-008-377-4

Query Match  
Best Local Similarity 83.0%; Score 39; DB 14; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 3 LAMIDLHV 10  
DB 1 LAMIDLHV 8

RESULT 4  
US-09-844-774-9  
Sequence 9, Application US/09844774  
Publication No. US20020037553A1  
GENERAL INFORMATION:  
APPLICANT: Al-Barazanj, Kamal A.  
APPLICANT: Arch, Jonathan Robert Sanders  
APPLICANT: Camilleri, Patrick  
APPLICANT: Neville, William Arthur  
TITLE OF INVENTION: Fragments of Leptin (OB protein)  
FILE REFERENCE: P31505-C1  
CURRENT APPLICATION NUMBER: US/09/844,774  
CURRENT FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: 09/154,866  
PRIOR FILING DATE: 1998-12-04  
PRIOR APPLICATION NUMBER: PCT/EP97/02968  
PRIOR FILING DATE: 1997-06-04  
PRIOR APPLICATION NUMBER: GB 9703493.8  
PRIOR FILING DATE: 1997-02-20  
PRIOR APPLICATION NUMBER: GB 9618540.0  
PRIOR FILING DATE: 1996-09-05  
PRIOR APPLICATION NUMBER: GB 9611775.9  
PRIOR FILING DATE: 1996-06-06  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 9  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-844-774-9

Query Match  
Best Local Similarity 55.3%; Score 26; DB 9; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 6 DLHV 10  
DB 1 DLHV 5

RESULT 5  
US-10-290-144-9  
Sequence 9, Application US/10290144  
Publication No. US20030092126A1  
GENERAL INFORMATION:  
APPLICANT: Al-Barazanj, Kamal A.  
APPLICANT: Arch, Jonathan Robert Sanders  
APPLICANT: Camilleri, Patrick  
APPLICANT: Neville, William Arthur  
TITLE OF INVENTION: Fragments of Leptin (OB protein)  
FILE REFERENCE: P31505C2  
CURRENT APPLICATION NUMBER: US/10/290,144  
CURRENT FILING DATE: 2002-11-07  
PRIOR APPLICATION NUMBER: 09/844,774  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: 09/194,866  
PRIOR FILING DATE: 1998-12-04  
PRIOR APPLICATION NUMBER: PCT/EP97/02968  
PRIOR FILING DATE: 1997-06-04  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 9  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-290-144-9

Query Match  
Best Local Similarity 55.3%; Score 26; DB 14; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 6 DLHV 10  
DB 1 DLHV 5

RESULT 6  
US-10-064-903-2  
Sequence 2, Application US/10064903  
Publication No. US20030059912A1  
GENERAL INFORMATION:  
APPLICANT: Biotechn Gesellschaft fur biotechnologische Entwicklung und Consulting  
APPLICANT: BMB  
TITLE OF INVENTION: HYBRID PROTEIN FOR INHIBITING THE DEGRADATION OF MASTOCYTES AND  
FILE REFERENCE: BIO-001PCT-CIP  
CURRENT APPLICATION NUMBER: US/10/064,903  
CURRENT FILING DATE: 2002-08-27  
PRIOR APPLICATION NUMBER: US 09/700,540  
PRIOR FILING DATE: 2001-01-19  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 8  
TYPE: PRT  
ORGANISM: Clostridium tectani  
US-10-064-903-2

Query Match 51.1%; Score 24; DB 14; Length 8;  
Best Local Similarity 80.0%; Pred. No. 1.4e+06;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 DLIHV 10  
DB 2 DLIHV 6

## RESULT 7

US-09-932-165-364  
; Sequence 364, Application US/09932165  
; Publication No. US20030134784A1  
; GENERAL INFORMATION:  
; APPLICANT: RAITANO, ARTHUR  
; APPLICANT: CHALLITA-EID, PIA M.  
; APPLICANT: FARIS, MARY  
; APPLICANT: SAFRAN, DOUGLAS  
; APPLICANT: AFAR, DANIEL  
; APPLICANT: LEVIN, ELANA  
; APPLICANT: HUBERT, RENE  
; APPLICANT: GE, WANGMAO  
; APPLICANT: JAKOBOWITS, AVA  
; TITLE OF INVENTION: NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED  
; TITLE OF INVENTION: 83P2H3 AND CATF2E11 USEFUL IN TREATMENT AND  
; FILE REFERENCE: 51158-20014.00  
; CURRENT APPLICATION NUMBER: US/09/932,165  
; CURRENT FILING DATE: 2001-08-17  
; PRIOR APPLICATION NUMBER: 60/226,329  
; PRIOR FILING DATE: 2000-08-17  
; NUMBER OF SEQ ID NOS: 1508  
; SOFTWARE: Patent Ver. 2.1  
; SEQ ID NO 364  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Peptide motif  
US-09-932-165-364

Query Match 51.1%; Score 24; DB 10; Length 10;  
Best Local Similarity 66.7%; Pred. No. 7.6e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LAMLDLIHV 10  
DB 2 LAALLLVHV 10

## RESULT 8

US-10-024-652-1548  
; Sequence 1548, Application US/10024652  
; Publication No. US20030219738A1  
; GENERAL INFORMATION:  
; APPLICANT: Agensys, Inc  
; APPLICANT: Challita-Eid, Pia M.  
; APPLICANT: Faris, Mary  
; APPLICANT: Afar, Daniel E.H.  
; APPLICANT: Hubert, Rene S.  
; APPLICANT: Mitchell, Steve Chappell  
; APPLICANT: Levin, Elana  
; APPLICANT: Morrison, Karen Jane Meyrick  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Jakobovits, Ava  
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc  
; TITLE OF INVENTION: Transporter Protein Entitled 108P5H8 Useful in Treatment and  
; TITLE OF INVENTION: Detection of Cancer  
; FILE REFERENCE: 51158-20025.00  
; CURRENT APPLICATION NUMBER: US/10/024,652  
; CURRENT FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: 60/256,210  
; PRIOR FILING DATE: 2000-12-15

NUMBER OF SEQ ID NOS: 2598  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1548  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: homo sapien

US-10-024-652-1548

Query Match 47.9%; Score 22.5; DB 14; Length 10;  
Best Local Similarity 70.0%; Pred. No. 1.4e+03;  
Matches 7; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 SLA-MDLIH 9  
DB 1 SLAINTDALH 10

## RESULT 9

US-10-024-652-1816  
; Sequence 1816, Application US/10024652  
; Publication No. US20030219738A1  
; GENERAL INFORMATION:  
; APPLICANT: Agensys, Inc  
; APPLICANT: Challita-Eid, Pia M.  
; APPLICANT: Faris, Mary  
; APPLICANT: Afar, Daniel E.H.  
; APPLICANT: Hubert, Rene S.  
; APPLICANT: Mitchell, Steve Chappell  
; APPLICANT: Levin, Elana  
; APPLICANT: Morrison, Karen Jane Meyrick  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Jakobovits, Ava  
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc  
; TITLE OF INVENTION: Transporter Protein Entitled 108P5H8 Useful in Treatment and  
; FILE REFERENCE: 51158-20025.00  
; CURRENT APPLICATION NUMBER: US/10/024,652  
; CURRENT FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: 60/256,210  
; PRIOR FILING DATE: 2000-12-15  
; NUMBER OF SEQ ID NOS: 2598  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1816  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: homo sapien

Query Match 47.9%; Score 22.5; DB 14; Length 10;  
Best Local Similarity 70.0%; Pred. No. 1.4e+03;  
Matches 7; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 SLA-MDLIH 9  
DB 1 SLAINTDALH 10

## RESULT 10

US-09-758-128-30  
; Sequence 30, Application US/09758128  
; Patent No. US20020107187A1  
; GENERAL INFORMATION:  
; APPLICANT: KINGSTON, David J.  
; APPLICANT: GERATY, No. US20020107187A1man L.  
; APPLICANT: WESTBROOK, Simon L.  
; TITLE OF INVENTION: MODULATING THE ACTIVITY OF HORMONES OR THEIR RECEPTORS  
; TITLE OF INVENTION: - PEPTIDES, ANTIBODIES, VACCINES AND USES THEREOF  
; FILE REFERENCE: 016786/0214  
; CURRENT APPLICATION NUMBER: US/09/758,128  
; CURRENT FILING DATE: 2001-01-12  
; PRIOR APPLICATION NUMBER: 09/194,218  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: AU PN9990

; PRIOR FILING DATE: 1996-05-22  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO: 30  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-758-128-30

Query Match 46.8%; Score 22; DB 9; Length 7;  
Best Local Similarity 60.0%; Pred. No. 1.4e+06;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 LDLLH 9  
:||||  
Db 1 MDMLH 5

RESULT 11  
US-09-758-426-30  
; Sequence 30, Application US/09758426  
; Patent No. US20020169116A1  
; GENERAL INFORMATION:  
; APPLICANT: KINGSTON, David J.  
; APPLICANT: GERRATY, No. US20020169116A1man L.  
; APPLICANT: WESTBROOK, Simon L.  
; TITLE OF INVENTION: MODULATING THE ACTIVITY OF HORMONES OR THEIR RECEPTORS  
; FILE REFERENCE: 016786/0214  
; CURRENT APPLICATION NUMBER: US/09/758,426  
; PRIOR FILING DATE: 2001-01-12  
; PRIOR APPLICATION NUMBER: 09/194,218  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: AU PN9990  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO: 30  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-758-426-30

Query Match 46.8%; Score 22; DB 9; Length 7;  
Best Local Similarity 60.0%; Pred. No. 1.4e+06;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 LDLLH 9  
:||||  
Db 1 MDMLH 5

RESULT 12  
US-09-758-198-30  
; Sequence 30, Application US/09758198  
; Publication No. US20020187925A1  
; GENERAL INFORMATION:  
; APPLICANT: KINGSTON, David J.  
; APPLICANT: GERRATY, No. US20020187925A1man L.  
; APPLICANT: WESTBROOK, Simon L.  
; TITLE OF INVENTION: MODULATING THE ACTIVITY OF HORMONES OR THEIR RECEPTORS  
; FILE REFERENCE: 016786/0214  
; CURRENT APPLICATION NUMBER: US/09/758,198  
; PRIOR FILING DATE: 2001-01-12  
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-02-05  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-22  
; PRIOR APPLICATION NUMBER: PCT/AU97/00312  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-05-22  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO: 30  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-758-198-30

Query Match 46.8%; Score 22; DB 9; Length 7;  
Best Local Similarity 60.0%; Pred. No. 1.4e+06;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 LDLLH 9  
:||||  
Db 1 MDMLH 5

RESULT 13  
US-09-861-661-30  
; Sequence 30, Application US/09861661  
; Publication No. US20030045676A1  
; GENERAL INFORMATION:  
; APPLICANT: KINGSTON, DAVID J.  
; APPLICANT: GERRATY, NORMAN L.  
; APPLICANT: WESTBROOK, SIMON L.  
; TITLE OF INVENTION: PEPTIDES, ANTIBODIES, VACCINES & USES THEREOF  
; FILE REFERENCE: 054270/0135  
; CURRENT APPLICATION NUMBER: US/09/861,661  
; PRIOR FILING DATE: 2001-05-22  
; PRIOR APPLICATION NUMBER: 09/194,218  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: AU PN9990  
; PRIOR FILING DATE: 1996-05-22  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO: 30  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-861-661-30

Query Match 46.8%; Score 22; DB 10; Length 7;  
Best Local Similarity 60.0%; Pred. No. 1.4e+06;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 LDLLH 9  
:||||  
Db 1 MDMLH 5

RESULT 14  
US-10-400-991-27  
; Sequence 27, Application US/10400991  
; Publication No. US20030224417A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Glucksmann, Maria A.  
; APPLICANT: Welch, Nadine S.  
; APPLICANT: Hunter, John J.  
; APPLICANT: White, David  
; APPLICANT: Macbeth, Kyle J.  
; APPLICANT: Williamson, Mark J.  
; APPLICANT: Meyers, Rachel E.  
; APPLICANT: Cunn, Mlyoung  
; TITLE OF INVENTION: 14400, 2838, 14618, 15334, 14274, 32164,  
; TITLE OF INVENTION: 39404, 38911, 26904, 31237, 18057, 16405, 32705, 23224,  
; TITLE OF INVENTION: 27423, 32700, 32712 AND 12216, NOVEL SEVEN-TRANSMEMBRANE  
; FILE REFERENCE: ME103-0240NIN  
; CURRENT APPLICATION NUMBER: US/10/400,991  
; PRIOR FILING DATE: 2003-03-27  
; PRIOR APPLICATION NUMBER: US 10/190,469  
; PRIOR FILING DATE: 2002-07-05  
; PRIOR APPLICATION NUMBER: US 09/439,159  
; PRIOR FILING DATE: 1999-11-12

```

; PRIOR APPLICATION NUMBER: US 09/137,063
; PRIOR FILING DATE: 1998-08-20
; PRIOR APPLICATION NUMBER: US 10/167,192
; PRIOR FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: US 09/420,187
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US 09/173,869
; PRIOR FILING DATE: 1998-10-16
; PRIOR APPLICATION NUMBER: US 10/339,056
; PRIOR FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: US 09/377,429
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: US 09/136,726
; PRIOR FILING DATE: 1998-08-18
; PRIOR APPLICATION NUMBER: US 09/911,583
; PRIOR FILING DATE: 2001-07-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Amino Acid Fragment
; US-10-400-991-27

```

```

Query Match          46.8%; Score 22; DB 14; Length 7;
Best Local Similarity 71.4%; Pred. No. 1.4e+06;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2 LAMDDL 8
        |||
Db      1 LAVADLL 7

```

```

RESULT 15
US-10-367-580-170
; Sequence 170, Application US/10367580
; Publication No. US20040071720A1
; GENERAL INFORMATION:
; APPLICANT: Rothman, James E.
; APPLICANT: Hartl, F. Ulrich
; APPLICANT: Hoe, Mee H.
; APPLICANT: Houghton, Alan
; APPLICANT: Takeuchi, Yoshizumi
; APPLICANT: Mayhew, Mark
; TITLE OF INVENTION: Heat Shock Protein-Based Vaccines and Immunotherapies
; FILE REFERENCE: 11746/461061
; CURRENT APPLICATION NUMBER: US/10/367,580
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: US 09/794,832
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: US 09/011,645
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: PCT/US96/13363
; PRIOR FILING DATE: 1996-08-16
; PRIOR APPLICATION NUMBER: US 60/002,490
; PRIOR FILING DATE: 1995-08-18
; PRIOR APPLICATION NUMBER: US 60/002,479
; PRIOR FILING DATE: 1995-08-18
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 170
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
; US-10-367-580-170

```

```

Query Match          46.8%; Score 22; DB 15; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;

```

```

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      6 DILL 9
        |||
Db      3 DILL 6

```

```

Search completed: November 30, 2004, 08:31:00
Job time : 45 secs

```

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2004, 07:41:57 ; Search time 27 Seconds

(without alignments)  
24.562 Million cell updates/sec

Title: US-10-008-377A-1

Perfect score: 47

Sequence: 1 SLAMDLHLV 10

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 110780

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	53.2	10	3	US-08-159-339A-649
2	23	48.9	10	4	US-09-601-729-268
3	22	46.8	9	4	US-09-393-634-88
4	21	44.7	7	3	US-09-258-754-114
5	21	44.7	7	3	US-09-047-107-414
6	21	44.7	7	4	US-09-722-250D-414
7	21	44.7	7	4	US-09-676-45A-414
8	21	44.7	8	2	US-08-917-320-13
9	21	44.7	8	1	US-08-917-320-13
10	21	44.7	9	1	PCT-US95-04611A-13
11	21	44.7	9	1	US-08-217-188A-3
12	21	44.7	9	1	US-08-687-226-3
13	21	44.7	9	3	US-08-687-226-3
14	21	44.7	9	3	US-09-007-748-3
15	21	44.7	9	3	US-09-053-003-16
16	21	44.7	9	3	US-09-183-931-31
17	21	44.7	9	3	US-09-705-160-31
18	21	44.7	9	4	US-08-197-484-81
19	21	44.7	9	4	US-08-197-484-140
20	21	44.7	9	4	US-08-818-094-16
21	21	44.7	9	5	PCT-US95-02121-81
22	21	44.7	10	1	US-08-217-188A-20
23	21	44.7	10	1	US-08-687-226-20
24	21	44.7	10	3	US-08-687-226B-20
25	21	44.7	10	3	US-09-007-748-20
26	21	44.7	10	4	US-08-197-484-82
27	21	44.7	10	4	US-08-197-484-141

28	21	44.7	10	4	US-08-980-326-18	Sequence 18, Appl
29	21	44.7	10	5	PCT-US95-02121-82	Sequence 82, Appl
30	21	44.7	10	5	PCT-US95-02121-141	Sequence 141, Appl
31	20	42.6	5	4	US-09-082-358B-10	Sequence 10, Appl
32	20	42.6	7	1	US-08-166-930-16	Sequence 16, Appl
33	20	42.6	7	2	US-08-727-045A-16	Sequence 16, Appl
34	20	42.6	7	3	US-09-223-587-1	Sequence 1, Appl
35	20	42.6	7	3	US-09-312-314C-1	Sequence 1, Appl
36	20	42.6	7	3	US-09-312-314C-1	Sequence 1, Appl
37	20	42.6	7	4	US-09-516-488A-8	Sequence 8, Appl
38	20	42.6	7	4	US-09-408-172-16	Sequence 16, Appl
39	20	42.6	7	4	US-09-394-234-8	Sequence 8, Appl
40	20	42.6	9	3	US-08-817-547A-13	Sequence 2, Appl
41	20	42.6	9	3	US-09-008-481A-13	Sequence 13, Appl
42	20	42.6	9	3	US-09-309-592-13	Sequence 13, Appl
43	20	42.6	10	1	US-08-843-035-17	Sequence 17, Appl
44	20	42.6	10	1	US-08-843-035-18	Sequence 18, Appl
45	20	42.6	10	1	US-08-843-035-19	Sequence 19, Appl

#### ALIGNMENTS

RESULT 1  
US-08-159-339A-649  
Sequence 649, Application US/08159339A  
Patent No. 6037135  
GENERAL INFORMATION:  
APPLICANT: Kubo, Ralph T.  
APPLICANT: Grey, Howard M.  
APPLICANT: Sette, Alessandro  
APPLICANT: Celis, Esteban  
TITLE OF INVENTION: HLA Binding peptides and their  
NUMBER OF SEQUENCES: 1254  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/159,339A  
FILING DATE: 29-NOV-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/926,666  
FILING DATE: 07-AUG-1992  
APPLICATION NUMBER: US 08/027,746  
FILING DATE: 05-MAR-1993  
APPLICATION NUMBER: US 08/103,396  
FILING DATE: 06-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Ellen Lawver  
REGISTRATION NUMBER: 32,762  
REFERENCE/DOCKET NUMBER: 018623-005030US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
TELEX:  
INFORMATION FOR SEQ ID NO: 649:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

US-08-159-339A-649

Query Match 53.2%; Score 25; DB 3; Length 10;  
 Best Local Similarity 44.4%; Pred. No. 1.6e+02;  
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 LAMDLHV 10  
 :|||:  
 DB 2 LQCVDLVHI 10

RESULT 2  
 US-09-601-729-268  
 ; Sequence 268; Application US/09601729  
 ; Patent No. 6683052  
 ; GENERAL INFORMATION:

APPLICANT: THIAM, KADER  
 APPLICANT: AURIAULT, CLAUDE  
 APPLICANT: GRAS-MASSE, HELENE  
 APPLICANT: LOING, ESTELLE  
 APPLICANT: VERWAERDE, CLAUDE  
 APPLICANT: GUILLET, JEAN GERARD  
 TITLE OF INVENTION: LIPOPEPTIDES CONTAINING AN INTERFERON FRAGMENT AND USES  
 TITLE OF INVENTION: THEREOF IN PHARMACEUTICAL COMPOSITIONS  
 FILE REFERENCE: US-97-AU-IN  
 CURRENT APPLICATION NUMBER: US/09/601,729  
 CURRENT FILING DATE: 2000-11-20  
 PRIOR APPLICATION NUMBER: PCT/FR99/00259  
 PRIOR FILING DATE: 1999-02-05  
 PRIOR APPLICATION NUMBER: 98 01439  
 PRIOR FILING DATE: 1998-02-06  
 NUMBER OF SEQ ID NOS: 281  
 SOFTWARE: Patent In Ver. 2.1  
 SEQ ID NO 268  
 LENGTH: 10  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 OTHER INFORMATION: peptide  
 US-09-601-729-268

Query Match 48.9%; Score 23; DB 4; Length 10;  
 Best Local Similarity 71.4%; Pred. No. 3.5e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AMLDLH 9  
 :|||:  
 DB 2 AALDLH 8

RESULT 3  
 US-09-393-634-88  
 ; Sequence 88; Application US/09393634  
 ; Patent No. 6558910  
 ; GENERAL INFORMATION:  
 APPLICANT: Zuker, Charles S.  
 APPLICANT: Adler, Jon Elliot  
 APPLICANT: Ryba, Nick  
 APPLICANT: Mueller, Ken  
 APPLICANT: Hoon, Mark  
 APPLICANT: The Regents of the University of California  
 APPLICANT: The Government of the United States of America  
 APPLICANT: as represented by the Secretary of the  
 APPLICANT: Department of Health and Human Services  
 TITLE OF INVENTION: SF, a No. 6558910e1 Family of Taste Receptors  
 FILE REFERENCE: 023075-098000US  
 CURRENT APPLICATION NUMBER: US/09/393,634  
 CURRENT FILING DATE: 1999-09-10  
 NUMBER OF SEQ ID NOS: 92  
 SOFTWARE: Patent In Ver. 2.1  
 SEQ ID NO 88  
 LENGTH: 9

TYPE: PRT  
 ORGANISM: Artificial Sequence

FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: SF01 signature  
 OTHER INFORMATION: sequence 1, amino acids encoded by PCR primers  
 OTHER INFORMATION: identifying polymorphic variants, interspecies  
 OTHER INFORMATION: homologs and alleles of SF family members  
 US-09-393-634-88

Query Match 46.8%; Score 22; DB 4; Length 9;  
 Best Local Similarity 71.4%; Pred. No. 3.8e+05;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAMDL 8  
 :|||:  
 DB 2 MAPDL 8

RESULT 4  
 US-09-258-754-414  
 ; Sequence 414; Application US/09258754  
 ; Patent No. 6174687  
 ; GENERAL INFORMATION:

APPLICANT: Ruostantti, Erkki  
 APPLICANT: Pasqualini, Renata  
 APPLICANT: Rajotte, Daniel  
 TITLE OF INVENTION: Methods of Identifying Lung Homing Molecules Using  
 FILE REFERENCE: P-LJ 3443  
 CURRENT APPLICATION NUMBER: US/09/258,754  
 CURRENT FILING DATE: 1999-02-26  
 EARLIER APPLICATION NUMBER: 09/042,107  
 EARLIER FILING DATE: 1998-03-13  
 NUMBER OF SEQ ID NOS: 452  
 SOFTWARE: Patent In Ver. 2.0  
 SEQ ID NO 414  
 LENGTH: 7  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 US-09-258-754-414

Query Match 44.7%; Score 21; DB 3; Length 7;  
 Best Local Similarity 83.3%; Pred. No. 3.8e+05;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMD 6  
 :|||:  
 DB 1 SLAMD 6

RESULT 5  
 US-09-042-107-414  
 ; Sequence 414; Application US/09042107  
 ; Patent No. 6232287  
 ; GENERAL INFORMATION:  
 APPLICANT: Ruostantti, Erkki  
 APPLICANT: Pasqualini, Renata  
 TITLE OF INVENTION: Molecules that Home to Various Selected Organs or  
 TITLE OF INVENTION: Tissues  
 FILE REFERENCE: P-LJ 2892  
 CURRENT APPLICATION NUMBER: US/09/042,107  
 CURRENT FILING DATE: 1998-03-13  
 NUMBER OF SEQ ID NOS: 436  
 SOFTWARE: Patent In Ver. 2.0  
 SEQ ID NO 414  
 LENGTH: 7  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 US-09-042-107-414

Tue Nov 30 08:49:15 2004

us-10-008-377a-1.closed.ral

Page 3

Query Match 44.7%; Score 21; DB 3; Length 7;  
Best Local Similarity 83.3%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMMD 6  
DB 1 SLAMMD 6

RESULT 6  
US-09-722-250D-414  
Sequence 414, Application US/09722250D  
Patent No. 6610651  
GENERAL INFORMATION:  
APPLICANT: Rucsiabhi, Erkki  
APPLICANT: Pasqualini, Renata  
TITLE OF INVENTION: Molecules that Home to Various Selected Organs or  
FILE REFERENCE: P-LJ 4514  
CURRENT APPLICATION NUMBER: US/09/722,250D  
CURRENT FILING DATE: 2000-11-22  
PRIOR APPLICATION NUMBER: US 09/042,107  
NUMBER OF SEQ ID NOS: 437  
SOFTWARE: Patent Ver. 2.0  
SEQ ID NO 414  
LENGTH: 7  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-722-250D-414

Query Match 44.7%; Score 21; DB 4; Length 7;  
Best Local Similarity 83.3%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMMD 6  
DB 1 SLAMMD 6

RESULT 7  
US-09-676-475A-414  
Sequence 414, Application US/09676475A  
Patent No. 6784153  
GENERAL INFORMATION:  
APPLICANT: Rucsiabhi, Erkki  
APPLICANT: Pasqualini, Renata  
APPLICANT: Rajotte, Daniel  
TITLE OF INVENTION: Methods of Identifying Lung Homing Molecules Using  
FILE REFERENCE: P-LA 4377  
CURRENT APPLICATION NUMBER: US/09/676,475A  
CURRENT FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 09/042,107  
NUMBER OF SEQ ID NOS: 452  
SOFTWARE: Patent Ver. 2.0  
SEQ ID NO 414  
LENGTH: 7  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-676-475A-414

Query Match 44.7%; Score 21; DB 4; Length 7;  
Best Local Similarity 83.3%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLAMMD 6

DB 1 SLAMMD 6

RESULT 8  
US-08-917-320-13  
Sequence 13, Application US/08917320  
Patent No. 5824508  
GENERAL INFORMATION:  
APPLICANT: Spaete, Richard and Jackman, Winthrop, T.  
TITLE OF INVENTION: No. 5824508 Splicing Variants of gp350/220  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooley Godward Castro Huddleson & Tatum  
STREET: 5 Palo Alto Square  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/917,320  
FILING DATE: 25-AUG-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/229,291  
FILING DATE: April 18, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Luanan Geary  
REGISTRATION NUMBER: 31,822  
REFERENCE/DOCKET NUMBER: AVIR-003/000US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-843-5163  
TELEFAX: 415-857-0663  
TELEX: 380816 CooleyPA  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
US-08-917-320-13

Query Match 44.7%; Score 21; DB 2; Length 8;  
Best Local Similarity 66.7%; Pred. No. 3.8e+05;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMMD 6  
DB 1 SLAMMD 6

RESULT 9  
PCT-US95-04611A-13  
Sequence 13, Application PC/TUS9504611A  
GENERAL INFORMATION:  
APPLICANT: Spaete, Richard and Jackman, Winthrop, T.  
TITLE OF INVENTION: Non Splicing Variants of gp350/220  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooley Godward Castro Huddleson & Tatum  
STREET: 5 Palo Alto Square  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04611A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/229,291  
FILING DATE: April 18, 1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Luanh Cseri  
REGISTRATION NUMBER: 31,822  
REFERENCE/DOCKET NUMBER: AVIR-003/00US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-843-5163  
TELEFAX: 415-857-0663  
TELEX: 380816 CooleyPA  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPES: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
PCT-US95-04611A-13

Query Match 44.7%; Score 21; DB 5; Length 8;  
Best Local Similarity 66.7%; Pred. No. 3.8e+05;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLAMD 6  
Db 1 NISMLD 6

RESULT 10  
US-08-217-188A-3  
Sequence 3, Application US/08217188A  
Patent No. 5554724  
GENERAL INFORMATION:  
APPLICANT: Melief, Cornelis J. M.  
APPLICANT: Visseren, M. J. W.  
APPLICANT: Kast, W. M.  
APPLICANT: van der Bruggen, Pierre  
APPLICANT: Boon-Falleur, Thierry  
TITLE OF INVENTION: Isolated Tumor Rejection Antigen  
TITLE OF INVENTION: Precursor MAGE-2 Derived Peptides, and Uses Thereof  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felle & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/217,188A  
FILING DATE: 24-MARCH-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 5554724man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5340  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acid residues  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-217-188A-3

Query Match 44.7%; Score 21; DB 1; Length 9;  
Best Local Similarity 50.0%; Pred. No. 3.8e+05;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 MLDLH 9  
Db 2 MVELVH 7

RESULT 11  
US-08-687-226-3  
Sequence 3, Application US/08687226  
Patent No. 5686068  
GENERAL INFORMATION:  
APPLICANT: Melief, Cornelis J. M.; Visseren, M. W.;  
APPLICANT: van der Burg, Sjoerd; van der Bruggen, Pierre;  
APPLICANT: Boon-Falleur, Thierry  
TITLE OF INVENTION: Isolated Peptides Derived From  
TITLE OF INVENTION: MAGE-2, Cytolytic T Cells Specific To Complexes Of  
TITLE OF INVENTION: Peptides And HLA-A2 Molecules, And Uses Thereof  
NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felle & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/687,226  
FILING DATE: 25-JULY-1996  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/217,188  
FILING DATE: 24-MARCH-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 5686068man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5447  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acid residues  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-687-226-3

Query Match 44.7%; Score 21; DB 1; Length 9;  
Best Local Similarity 50.0%; Pred. No. 3.8e+05;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 MLDLH 9  
Db 2 MVELVH 7

RESULT 12  
US-08-667-725B-3  
Sequence 3, Application US/08667725B

Tue Nov 30 08:49:15 2004

us-10-008-377a-1.closed.rat

Page 5

```
/ Patent No. 6063900
/ GENERAL INFORMATION:
/ APPLICANT: Melief, Cornelis J. M.
/ APPLICANT: Visseren, M. J. W.
/ APPLICANT: Kast, W. M.
/ APPLICANT: van der Bruggen, Pierre
/ APPLICANT: Boon-Falleur, Thierry
/ TITLE OF INVENTION: Isolated Tumor Rejection Antigen
/ TITLE OF INVENTION: Precursor MAGE-2 Derived Peptides, and Uses Thereof
/ NUMBER OF SEQUENCES: 64
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Fulbright & Jaworski LLP
/ STREET: 666 Fifth Avenue
/ CITY: New York City
/ STATE: New York
/ COUNTRY: USA
/ ZIP: 10103
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
/ OPERATING SYSTEM: PC-DOS
/ SOFTWARE: Wordperfect
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/667,725B
/ FILING DATE: 21 June 1996
/ CLASSIFICATION: 530
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Hanson, No. 6063900man D.
/ REGISTRATION NUMBER: 30,946
/ REFERENCE/DOCKET NUMBER: LUD 5340.1 DIV (081585)
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 318-3000
/ TELEFAX: (212) 752-5558
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 9 amino acid residues
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-667-725B-3

Query Match 44.7%; Score 21; DB 3; Length 9;
Best Local Similarity 50.0%; Pred. No. 3.de+05;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 MDLH 9
Db 2 MWELVH 7

RESULT 13
US-09-007-748-3
/ Sequence 3, Application US/09007748
/ Patent No. 6147187
/ GENERAL INFORMATION:
/ APPLICANT: Melief, Cornelis J. M.
/ APPLICANT: Visseren, M. J. W.
/ APPLICANT: Kast, W. M.
/ APPLICANT: van der Bruggen, Pierre
/ APPLICANT: Boon-Falleur, Thierry
/ TITLE OF INVENTION: Isolated Tumor Rejection Antigen
/ TITLE OF INVENTION: Precursor MAGE-2 Derived Peptides, and Uses Thereof
/ NUMBER OF SEQUENCES: 64
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Fulbright & Jaworski LLP
/ STREET: 666 Fifth Avenue
/ CITY: New York City
/ STATE: New York
/ COUNTRY: USA
/ ZIP: 10103
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
/ COMPUTER: IBM PS/2
```

```
/ OPERATING SYSTEM: PC-DOS
/ SOFTWARE: Wordperfect
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/007,748
/ FILING DATE: 15 January 1998
/ CLASSIFICATION: 530
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Hanson, No. 6147187man D.
/ REGISTRATION NUMBER: 30,946
/ REFERENCE/DOCKET NUMBER: LUD 5340.2 DIV (081572)
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 318-3000
/ TELEFAX: (212) 752-5558
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 9 amino acid residues
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-09-007-748-3

Query Match 44.7%; Score 21; DB 3; Length 9;
Best Local Similarity 50.0%; Pred. No. 3.de+05;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 MDLH 9
Db 2 MWELVH 7

RESULT 14
US-09-053-003-16
/ Sequence 16, Application US/09053003
/ Patent No. 6207391
/ GENERAL INFORMATION:
/ APPLICANT: Wu, Pengguang
/ APPLICANT: McKinney, Judi
/ TITLE OF INVENTION: High-Throughput Screening Assays for
/ TITLE OF INVENTION: Modulators of STAT4 and STAT6 Activity
/ NUMBER OF SEQUENCES: 64
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Townsend and Townsend and Crew LLP
/ STREET: Two Embarcadero Center, Eighth floor
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94111-3834
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/053,003
/ FILING DATE: 31-MAR-1998
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Smith, Timothy L.
/ REGISTRATION NUMBER: 35,367
/ REFERENCE/DOCKET NUMBER: 018781-0008000US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 576-0200
/ TELEFAX: (415) 576-0300
/ INFORMATION FOR SEQ ID NO: 16:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 9 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FEATURE:
/ NAME/KEY: Modified-site
/ LOCATION: 2
```

OTHER INFORMATION: /product= "OTHER"  
 OTHER INFORMATION: /note= "Xaa = Tyr or phosphotyrosine"  
 US-09-053-003-16

Query Match 44.7%; Score 21; DB 3; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 3.8e+05;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 LDLLHY 10  
 ||:|  
 Db 3 LDMPHV 8

RESULT 15  
 US-09-183-931-31  
 ; Sequence 31, Application US/09183931C

; Patent No. 6210886

; GENERAL INFORMATION:

; APPLICANT: Van Baren, Nicolas

; APPLICANT: Brasseur, Francis

; APPLICANT: Boon-Falleur, Thierry

; TITLE OF INVENTION: METHOD FOR DIAGNOSING LEUKEMIA BY DETERMINING

; FILE REFERENCE: LUD 5527.1-JEU/ES

; CURRENT APPLICATION NUMBER: US/09/183,931C

; CURRENT FILING DATE: 2000-02-28

; EARLIER APPLICATION NUMBER: US 09/018,422

; EARLIER FILING DATE: 1998 - 02 - 04

; NUMBER OF SEQ ID NOS: 44

; SEQ ID NO 31

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION:

; US-09-183-931-31

Query Match 44.7%; Score 21; DB 3; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 3.8e+05;  
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 MLDLH 5  
 ||:|  
 Db 2 MVELVH 7

Search completed: November 30, 2004, 08:15:58  
 Job time : 28 secs